

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2003, 12:40:24 (Search time 71 seconds
(without alignments)
26.999 Million cell updates/sec

Title: US-09-806-302A-2

Perfect score: 496

Sequence: 1 MKLWVLMIAALLHCYADS.....NFGIMHTYDSIMCKSN 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database:

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	456	100.0	95	10	US-09-110-716-31
2	496	100.0	95	10	US-09-985-911-6
3	408	82.3	77	10	US-09-110-716-13
4	393.5	79.3	76	10	US-09-110-716-40
5	281	56.7	93	10	US-09-757-417-27
6	281	56.7	93	10	US-09-934-054-3
7	281	56.7	93	10	US-09-934-054-10
8	281	56.7	93	10	US-09-934-054-10
9	281	56.7	132	10	US-09-757-417-47
10	281	56.7	410	12	US-10-007-805-495
11	281	56.7	743	12	US-10-007-805-494
12	281	56.7	1095	12	US-10-007-805-493
13	206	41.5	75	10	US-09-110-716-41
14	196.5	39.6	95	10	US-09-985-911-27
15	195.5	39.4	95	10	US-09-934-054-12
16	166	33.5	76	10	US-09-110-716-39
17	77	15.5	21	10	US-09-757-417-3
18	77	15.5	21	10	US-09-757-417-16
19	77	15.5	21	10	US-09-757-417-18

20	77	15.5	21	12	US-10-007-805-498	Sequence 498, App
21	72	14.5	20	10	US-09-757-417-32	Sequence 32, Appl
22	68	13.7	55	10	US-09-757-417-28	Sequence 28, Appl
23	67.5	13.6	112	10	US-09-934-054-8	Sequence 8, Appl
24	67	13.5	20	10	US-09-757-417-30	Sequence 30, Appl
25	66	13.3	20	10	US-09-757-417-33	Sequence 43, Appl
26	66	13.3	20	10	US-09-757-417-33	Sequence 33, Appl
27	66	13.3	20	12	US-10-007-805-499	Sequence 499, App
28	65.5	13.2	142	10	US-09-821-839-7	Sequence 9, Appl
29	62.5	12.6	91	10	US-09-934-054-9	Sequence 9, Appl
30	62.5	12.6	143	10	US-09-925-525-42	Sequence 42, Appl
31	62.5	12.6	293	9	US-09-925-525-42	Sequence 42, Appl
32	61	12.3	765	10	US-09-934-054-4	Sequence 4, Appl
33	61	12.3	766	9	US-09-934-054-4	Sequence 2, Appl
34	61	12.3	766	10	US-09-934-054-4	Sequence 2, Appl
35	60.5	12.2	180	9	US-09-985-911-256	Sequence 256, App
36	60.5	12.2	180	9	US-09-989-256-256	Sequence 256, App
37	60.5	12.2	180	9	US-09-989-735-256	Sequence 256, App
38	60.5	12.2	180	9	US-09-989-735-256	Sequence 256, App
39	60.5	12.2	180	9	US-09-989-730-256	Sequence 256, App
40	60.5	12.2	180	9	US-09-989-730-256	Sequence 256, App
41	60.5	12.2	180	9	US-09-989-730-256	Sequence 256, App
42	60.5	12.2	180	9	US-09-989-730-256	Sequence 256, App
43	60.5	12.2	180	9	US-09-989-734-256	Sequence 256, App
44	60.5	12.2	180	9	US-09-989-734-256	Sequence 256, App
45	60.5	12.2	180	9	US-10-174-590-196	Sequence 196, App

ALIGNMENTS

```
RESULT 1
US-09-110-716-31
Sequence 31, Application US/09110716A
Patent No. US20020034739A1
GENERAL INFORMATION:
APPLICANT: Lebrat, Robert I.
APPLICANT: Zhao, Chengquan
TITLE OF INVENTION: PEPTIDES CHARACTERISTIC OF CERTAIN TUMORS
FILE REFERENCE: 22000-20596.00
CURRENT APPLICATION NUMBER: US/09/110.716A
CURRENT FILING DATE: 1998-07-07
NUMBER OF SEQ ID NOS: 41
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 31
LENGTH: 95
TYPE: PPT
ORGANISM: Homo sapiens
US-09-110-716-31

Query Match          100.0%: Score 496; DB 10; Length 95;
Best Local Similarity 100.0%: Pred. No. 4.5e-48;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLWVLMIAALLHCYADSGCKLLEDEYKTEKLSIFRYELLQERFDSQAAEANG 60
DB 1 MKLWVLMIAALLHCYADSGCKLLEDEYKTEKLSIFRYELLQERFDSQAAEANG 60
QY 61 KFKCPTLNSHRTLNFGIMHTYDSIMCKSN 95
DB 61 KFKCPTLNSHRTLNFGIMHTYDSIMCKSN 95

RESULT 2
US-09-985-911-6
Sequence 6, Application US/09985911
Patent No. US20020151012A1
GENERAL INFORMATION:
APPLICANT: NI ET AL.
TITLE OF INVENTION: HUMAN ENDOMETRIAL SPECIFIC STEROID-BINDING FACTOR I, II AND III
FILE REFERENCE: PFS57D3
CURRENT APPLICATION NUMBER: US/09/985.911
```

THIS PAGE BLANK (use 70)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2003, 12:40:24 (Search time 71 seconds
(without alignments)
26.999 Million cell updates/sec)

Title: US-09-806-302a-2
Perfect score: 496
Sequence: 1 MKLWVLMALALHLCYADS.....NEGLMHTVDSIMCNMKN 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	496	100.0	95	10	US-09-110-716-31
2	496	100.0	95	10	US-09-985-911-6
3	408	82.3	77	10	US-09-110-716-13
4	393.5	79.3	76	10	US-09-110-716-40
5	281	56.7	93	10	US-09-757-417-27
6	281	56.7	93	10	US-09-934-054-3
7	281	56.7	93	10	US-09-934-054-10
8	281	56.7	93	12	US-10-007-805-503
9	281	56.7	132	10	US-09-757-417-47
10	281	56.7	140	12	US-10-007-805-495
11	281	56.7	743	12	US-10-007-805-494
12	281	56.7	1095	12	US-10-007-805-493
13	206	41.5	75	10	US-09-110-716-41
14	196.5	39.6	95	10	US-09-985-911-27
15	195.5	39.4	95	10	US-09-934-054-12
16	166	33.5	76	10	US-09-110-716-39
17	77	15.5	21	10	US-09-757-417-3
18	77	15.5	21	10	US-09-757-417-16
19	77	15.5	21	10	US-09-757-417-18

20	77	15.5	21	12	US-10-007-805-498	Sequence 498, App
21	72	14.5	20	10	US-09-757-417-32	Sequence 32, Appl
22	68	13.7	55	10	US-09-757-417-28	Sequence 28, Appl
23	67.5	13.6	112	10	US-09-934-054-8	Sequence 8, Appl
24	67	13.5	20	10	US-09-757-417-30	Sequence 30, Appl
25	66	13.3	20	10	US-09-757-417-4	Sequence 4, Appl
26	66	13.3	20	10	US-09-757-417-43	Sequence 43, Appl
27	66	13.3	20	12	US-10-007-805-499	Sequence 499, App
28	65.5	13.2	142	10	US-09-821-839-7	Sequence 7, Appl
29	62.5	12.6	91	10	US-09-934-054-9	Sequence 9, Appl
30	62.5	12.6	143	10	US-09-821-839-6	Sequence 6, Appl
31	62.5	12.6	293	3	US-09-902-525-42	Sequence 42, Appl
32	61	12.3	765	10	US-09-975-326-4	Sequence 4, Appl
33	61	12.3	766	10	US-09-934-406-2	Sequence 2, Appl
34	61	12.3	766	10	US-09-925-326-2	Sequence 2, Appl
35	60.5	12.2	180	9	US-09-992-598-256	Sequence 256, App
36	60.5	12.2	180	9	US-09-989-2928-256	Sequence 256, App
37	60.5	12.2	180	9	US-09-989-733-256	Sequence 256, App
38	60.5	12.2	180	9	US-09-990-444-256	Sequence 256, App
39	60.5	12.2	180	9	US-09-989-730-256	Sequence 256, App
40	60.5	12.2	180	9	US-09-990-436-256	Sequence 256, App
41	60.5	12.2	180	9	US-09-991-181-256	Sequence 256, App
42	60.5	12.2	180	9	US-09-993-687-256	Sequence 256, App
43	60.5	12.2	180	9	US-09-989-734-256	Sequence 256, App
44	60.5	12.2	180	9	US-09-997-653-256	Sequence 256, App
45	60.5	12.2	180	9	US-10-174-590-196	Sequence 196, App

ALIGNMENTS

```
RESULT 1
US-09-110-716-31
: Sequence 31, Application US/09110716A
: Patent No. US20020034739A1
: GENERAL INFORMATION:
: APPLICANT: Lehner, Robert I.
: APPLICANT: Zhao, Chengquan
: APPLICANT: Glasgow, Benjamin J.
: TITLE OF INVENTION: PEPTIDES CHARACTERISTIC OF CERTAIN TUMORS
: FILE REFERENCE: 22000-20596.00
: CURRENT APPLICATION NUMBER: US/09/110, 716A
: CURRENT FILING DATE: 1998-07-07
: NUMBER OF SEQ ID NOS: 41
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 31
: LENGTH: 95
: TYPE: PPT
: ORGANISM: lipophilin C
US-09-110-716-31

Query Match          100.0%; Score 496; DB 10; Length 95;
Best Local Similarity 100.0%; Pred. No. 4.5e-48;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLWVLMALALHLCYADSGCKLDEPVEKTSNDSIPREKELQFIDSDAAAFAMG 60
DB 1 MKLWVLMALALHLCYADSGCKLDEPVEKTSNDSIPREKELQFIDSDAAAFAMG 60
OY 61 KKKCFQNSHRTLNKFGIMHTVDSIMCNMKN 95
DB 61 KKKCFQNSHRTLNKFGIMHTVDSIMCNMKN 95

RESULT 2
US-09-985-911-6
: Sequence 6, Application US/09985911
: Patent No. US20020151012A1
: GENERAL INFORMATION:
: APPLICANT: NI ET AL.
: TITLE OF INVENTION: HUMAN ENDOMETRIAL SPECIFIC STEROID-BINDING FACTOR I, II AND III
: FILE REFERENCE: PF25703
: CURRENT APPLICATION NUMBER: US/09/985,911
```

```

?       CURRENT FILING DATE:   2001-11-06
? /      PRIOR APPLICATION NUMBER: 09/5883,1659
? /      PRIOR FILING DATE:    2000-05-30
? /      PRIOR APPLICATION NUMBER: 09/263,810
? /      PRIOR FILING DATE:    1999-03-08
? /      PRIOR APPLICATION NUMBER: 08/821,451
? /      PRIOR FILING DATE:    1997-03-21
? /      PRIOR APPLICATION NUMBER: 60/014,724
? /      PRIOR FILING DATE:    1996-03-21
? /      NUMBER OF SEQ ID NOS: 27
? /      SOFTWARE: PatentIn version 3.1.1
? /      SEQ ID NO 6
? /      LENGTH: 95
? /      TYPE: prt
? /      ORGANISM: human
? /      OS-09-985-911-6

```

Query Match	100.0%	Score 496;	DB 10;	Length 95;
Best Local Similarity	100.0%	Pred. No. 4.5e-48;		
Matches 95; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 3
US-09-110-716-13
? Sequence 13, Application US/09110716A
? Patent No. US20020034739A1
? GENERAL INFORMATION:
? APPLICANT: Lehrer, Robert I.
? APPLICANT: zhaoc, Chengquan
? APPLICANT: Glasgow, Benjamin J.
? TITLE OF INVENTION: PEPTIDES CHARACTERISTIC OF CERTAIN TUMORS
? FILE REFERENCE: 22000-20596.00
? CURRENT FILING DATE: US/09/110,716A
? NUMBER OF SEQ ID NOS: 41
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 13
? LENGTH: 77
? TYPE: PRT
? ORGANISM: Homo sapiens
? US-09-110-716-13

```

Query Match	82.38;	Score 408;	DB 10;	Length 77;
Best Local Similarity	100.08;	Pred. NO. 1.8e-38;		
Matches 77;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Oy 19 DSGCALLEDMVEKTIINDISIPEYKELLQEFIDSDAAAEAMGFKQCFINQSHRLTKNFG 78
|||||
|||
Db 1 DSGCKLLIEDMVEKTIINDISIPEYKELQEFIDSDAAAEAMGFKQCFINQSHRLTKNFG 600

```
QY      79 LMMHTVYDSIWCNMSN 95
        ||| ||| ||| ||| |||
Db      61 LMMHTVYDSIWCNMSN 77
```

RESULT 4 -
US-09-110-716-40
Sequence 40, Application US/09110716A
Patent No. US20020034739A1
GENERAL INFORMATION:
APPLICANT: Lehrer, Robert I.
APPLICANT: Zhao, Chengquan
APPLICANT: Glasgow, Benjamin J.
TITLE OF INVENTION: PEPTIDES CHARACTERISTIC OF CERTAIN TUMORS
FILE REFERENCE: 22000-70596 .00

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Query Match	79.38;	Score 393.5;	DB 10;	Length 76;
Best Local Similarity	98.78;	Pred. No. 7.1e-37;		
Matches 76;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1;

QY	LMHTYYDSTWCMKNSN	95
QY	19 DSGCKLLEEDVETKTSIDPSIPEYKELLQEFETDSDAAAEANGKRCQCFLNQSHRTLTAKNG	78
Db	1 DSGCKLLEEDVETKTSIDPSIPEYKELLQEFETDSDAAAEANGKRCQCFLNQSHRTLTAKNG	600
QY	79 LMHTYYDSTWCMKNSN	95
Db	61 LMHTT-YDSTWCMKNSN	76

US-09-757-417-27
Sequence 27, Application US/09757417
Patent No. US20020082216A1
GENERAL INFORMATION:
APPLICANT: Fanger, Gary R.
APPLICANT: Foy, Theresa M.
APPLICANT: Houghton, Raymond L.
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TITLE OF INVENTION: THERAPY, DIAGNOSIS AND MONITORING OF BREAST CANCER
FILE REFERENCE: 210121.479C
CURRENT APPLICATION NUMBER: US/09/757,417
CURRENT FILING DATE: 2001-01-08
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 27
LENGTH: 93
TYPE: PRT
ORGANISM: Homo sapien
US-09-757-417-27

Query Match	56.78;	Score 281;	DB 10;	Length 93;
Best Local Similarity	58.78;	pred. No. 2.3e-24;		
Matches	54;	Conservative	13;	Mismatches 25;
			Indels	0;
			Gaps	0;

```
OY      1 MLLVLMALILHCYADSGCKLIEDNVEKTIINSDISIPEKKELQEIIDSPAAREMG 60
        |||||  |||  |||  |||  :  |||  |||  :  |
Db      1 MKLVMLMALISQHCVAGSGCPLEENIVSKTINPQVSKEYKELLQEIIDNNATNAID 60
```

```
QY      61 KEKQCFNLNQSHTLKNFGLMNHVYDSIWCNM 92  
       : -|::|||: || - : -|::||| -::  
Db      61 ELKECFNLQTDETLISNEVEFMQLIYDSSLCDL 92
```

RESULT 6
US-09-93A-05A-3
Sequence 3, Application US/0993A054
Patent No. US20020107385A1
GENERAL INFORMATION:
APPLICANT: Akribiom, Inc/rid E.
Hillman, Jennifer L.
Murry, Lynn E.
Goli, Sura K.
Hawkins, Phillip R.
TITLE OF INVENTION: BREAST TUMOR SPECIFIC PROTEINS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 2200 Sand Hill Road, Suite 100

CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025-6936
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/934,054
FILING DATE: 21-Aug-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/747,547
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0077 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: <Unknown>
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-934-054-3

Query Match 56.7%; Score 281; DB 10; Length 93;
Best Local Similarity 58.7%; Pred. No. 2,3e-24;

Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

QY 1 KMLVWVLMALALHCHYADSGCKLEDMVEKTNISDIPIPEYKELLOEFIDSDAAAEAMG 60
DB 1 KMLVWVLMALALSOHCYAGSGCPLEENVISKTNIPVSKTEYKELLOEFIDSNATTNID 60
QY 61 KFKOCFLNOSHRTLKNGFLMHTVYDSIWCNM 92
DB 61 ELKECFLNQTDETLSNVEVFMOQLIYDSSLCDL 92
RESULT 7
US-09-934-054-10
Sequence 10, Application US/09934054
Patent No. US20020107385A1
GENERAL INFORMATION:
APPLICANT: Akerbom, Ingrid E.
Hillman, Jennifer L.
Murry, Lynn E.
Goli, Surya K.
Hawkins, Phillip R.
TITLE OF INVENTION: BREAST TUMOR SPECIFIC PROTEINS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025-6936
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/934,054
FILING DATE: 21-Aug-2001
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/747,547
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0077 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: <Unknown>
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-934-054-10

Query Match 56.7%; Score 281; DB 10; Length 93;
Best Local Similarity 58.7%; Pred. No. 2,3e-24;

Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

QY 1 KMLVWVLMALALHCHYADSGCKLEDMVEKTNISDIPIPEYKELLOEFIDSDAAAEAMG 60
DB 1 KMLVWVLMALALSOHCYAGSGCPLEENVISKTNIPVSKTEYKELLOEFIDSNATTNID 60
QY 61 KFKOCFLNOSHRTLKNGFLMHTVYDSIWCNM 92
DB 61 ELKECFLNQTDETLSNVEVFMOQLIYDSSLCDL 92

RESULT 8
US-10-007-805-503
Sequence 503, Application US/10007805
Patent No. US20020150581A1
GENERAL INFORMATION:

APPLICANT: Jiang, Yugu
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: McNeill, Patricia D.
APPLICANT: Durham, Margarita
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.470C10
CURRENT APPLICATION NUMBER: US/10/007,805
CURRENT FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 593
SOFTWARE: PastsEQ for Windows Version 4.0
SEQ ID NO 503
LENGTH: 93
TYPE: PRT
ORGANISM: Homo sapiens
US-10-007-805-503

Query Match 56.7%; Score 281; DB 12; Length 93;
Best Local Similarity 58.7%; Pred. No. 2,3e-24;

Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

QY 1 KMLVWVLMALALHCHYADSGCKLEDMVEKTNISDIPIPEYKELLOEFIDSDAAAEAMG 60
DB 1 KMLVWVLMALALSOHCYAGSGCPLEENVISKTNIPVSKTEYKELLOEFIDSNATTNID 60
QY 61 KFKOCFLNOSHRTLKNGFLMHTVYDSIWCNM 92
DB 61 ELKECFLNQTDETLSNVEVFMOQLIYDSSLCDL 92

RESULT 9
US-09-757-417-47
Sequence 47, Application US/09757417
Patent No. US20020082216A1
GENERAL INFORMATION:
APPLICANT: Fanger, Gary R.
APPLICANT: Foy, Theresa M.
APPLICANT: Houghton, Raymond L.
APPLICANT: Reed, Steven G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
THERAPY, DIAGNOSIS AND MONITORING OF BREAST CANCER
FILE REFERENCE: 210121.479C1
CURRENT APPLICATION NUMBER: US/09/757,417
CURRENT FILING DATE: 2001-01-08
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 47
LENGTH: 132
TYPE: PRT
ORGANISM: Homo sapiens
US-09-757-417-47

Query Match 56.7%: Score 281; DB 10; Length 132;
Best Local Similarity 58.7%: Pred. No. 3, 6e-24;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

QY 1 MKLWVLMALALLHCTADSGCKLEDMVEKTTNSDISPEYKELLOEFTSDAAAEAMG 60
DB 40 MKLWVLMALALSOHCYAGSCGPLEENVISKTTINPOVSKTEYKELLOEFTIDNATTNAID 99
DB 61 KFKOCFLNOSHRTLKNFGMLMHTVYDSIMCNCM 92
DB 100 ELKECFLNQTDETLSNVEFMQLIYDSSLCDL 131

RESULT 10
US-10-007-805-495
Sequence 495, Application US/10007805
Patent No. US20020150581A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: McNeill, Patricia D.
APPLICANT: Durham, Margarita
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITL OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470C10
CURRENT APPLICATION NUMBER: US/10/007,805
CURRENT FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 593
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 495
LENGTH: 410
TYPE: PRT
ORGANISM: Homo sapiens
US-10-007-805-495

Query Match 56.7%: Score 281; DB 12; Length 410;
Best Local Similarity 58.7%: Pred. No. 1, 4e-23;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

QY 1 MKLWVLMALALLHCTADSGCKLEDMVEKTTNSDISPEYKELLOEFTSDAAAEAMG 60
DB 1 MKLWVLMALALSOHCYAGSCGPLEENVISKTTINPOVSKTEYKELLOEFTIDNATTNAID 60

QY 61 KFKOCFLNOSHRTLKNFGMLMHTVYDSIMCNCM 92
DB 61 ELKECFLNQTDETLSNVEFMQLIYDSSLCDL 92

RESULT 11
US-10-007-805-494
Sequence 494, Application US/10007805
Patent No. US20020150581A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: McNeill, Patricia D.
APPLICANT: Durham, Margarita
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITL OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470C10
CURRENT APPLICATION NUMBER: US/10/007,805
CURRENT FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 593
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 494
LENGTH: 743
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: 403, 322, 615
OTHER INFORMATION: Xaa - Any Amino Acid
US-10-007-805-494

Query Match 56.7%: Score 281; DB 12; Length 743;
Best Local Similarity 58.7%: Pred. No. 2, 8e-23;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

QY 1 MKLWVLMALALLHCTADSGCKLEDMVEKTTNSDISPEYKELLOEFTSDAAAEAMG 60
DB 1 MKLWVLMALALSOHCYAGSCGPLEENVISKTTINPOVSKTEYKELLOEFTIDNATTNAID 60
QY 61 KFKOCFLNOSHRTLKNFGMLMHTVYDSIMCNCM 92
DB 61 ELKECFLNQTDETLSNVEFMQLIYDSSLCDL 92

RESULT 12
US-10-007-805-493
Sequence 493, Application US/10007805
Patent No. US20020150581A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Xu, Jiangchun
APPLICANT: Harlocker, Susan L.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: McNeill, Patricia D.
APPLICANT: Durham, Margarita
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITL OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.470C10
CURRENT APPLICATION NUMBER: US/10/007,805
CURRENT FILING DATE: 2001-12-07
NUMBER OF SEQ ID NOS: 593

```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 493
; LENGTH: 1095
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 403, 522, 615
; OTHER INFORMATION: xaa - Any Amino Acid
US-10-007-805-493

Query Match
Best Local Similarity 56.7%; Score 281; DB 12; Length 1095;
Best Local Similarity 58.7%; Pred. No. 4,4e-23;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

QY 1 MRLVLAALLLHCVADSGCKLDEWVEKTIKNSDISPEKKELQEFIDSDAAAEAMG 60
DB 1 MRLVLAALLLHCVADSGCKLDEWVEKTIKNSDISPEKKELQEFIDSDAAAEAMG 60
QY 61 KFKQCFLNQSHRTLNKFGIMMHTVYDSINCKM 92
DB 61 ELKECFLNQSHRTLNKFGIMMHTVYDSINCKM 92

RESULT 13
US-09-110-716-41
; Sequence 41, Application US/09110716A
; Patent No. US20020034739A1
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Zhao, Chengquan
; APPLICANT: Glasgow, Benjamin J.
; TITLE OF INVENTION: PEPTIDES CHARACTERISTIC OF CERTAIN TUMORS
; FILE REFERENCE: 22000-20596.00
; CURRENT APPLICATION NUMBER: US/09/110,716A
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Mmg
US-09-110-716-41

Query Match
Best Local Similarity 41.5%; Score 206; DB 10; Length 75;
Best Local Similarity 52.1%; Pred. No. 3,4e-16;
Matches 38; Conservative 13; Mismatches 22; Indels 0; Gaps 0;

QY 20 SGCKLDEWVEKTIKNSDISPEKKELQEFIDSDAAAEAMGKFKQCFLNQSHRTLNKFG 79
DB 2 SGCKLDEWVEKTIKNSDISPEKKELQEFIDSDAAAEAMGKFKQCFLNQSHRTLNKFG 79
QY 80 MMHTVYDSINCKM 92
DB 62 FMQLYDSSLCDL 74

RESULT 14
US-09-985-911-27
; Sequence 27, Application US/09985911
; Patent No. US20020151012A1
; GENERAL INFORMATION:
; APPLICANT: NI ET AL.
; TITLE OF INVENTION: HUMAN ENDOMETRIAL SPECIFIC STEROID-BINDING FACTOR I, II AND III
; FILE REFERENCE: PF257D3
; CURRENT APPLICATION NUMBER: US/09/985,911
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 09/383,169
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 09/263,810
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 08/821,451
; PRIOR FILING DATE: 1997-03-21
```

```
; PRIOR APPLICATION NUMBER: 60/014,724
; PRIOR FILING DATE: 1996-03-21
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
; LENGTH: 95
; TYPE: PRT
; ORGANISM: human
US-09-985-911-27

Query Match
Best Local Similarity 39.6%; Score 196.5; DB 10; Length 95;
Best Local Similarity 35.8%; Pred. No. 5,1e-15;
Matches 34; Conservative 31; Mismatches 29; Indels 1; Gaps 1;

QY 1 MRLVLAALLLHCVADSGCKLDEWVEKTIKNSDISPEKKELQEFIDSDAAAEAM 59
DB 1 MRLVLAALLLHCVADSGCKLDEWVEKTIKNSDISPEKKELQEFIDSDAAAEAM 59
QY 60 GFKQCFLNQSHRTLNKFGIMMHTVYDSINCKMKS 94
DB 61 KFKQCFLNQSHRTLNKFGIMMHTVYDSINCKMKS 94

RESULT 15
US-09-934-054-12
; Sequence 12, Application US/09934054
; Patent No. US20020107385A1
; GENERAL INFORMATION:
; APPLICANT: Akervold, Ingrid E.
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; APPLICANT: Murty, Lynn E.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: BREAST TUMOR SPECIFIC PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 2200 Sand Hill Road, Suite 100
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025-6936
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/934,054
; FILING DATE: 21-Aug-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,547
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0077 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: <Unknown>
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-934-054-12

Query Match
Best Local Similarity 39.4%; Score 195.5; DB 10; Length 95;
Best Local Similarity 35.8%; Pred. No. 6,6e-15;
Matches 34; Conservative 30; Mismatches 30; Indels 1; Gaps 1;
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2003, 13:19:09 (Search time 82 Seconds
(without alignments)
48,052 Million cell updates/sec

Title: US-09-806-302a-2
Perfect score: 496
Sequence: 1 MKLLMVLMLAALLHICVADS.....NEGIMHTVYDSIWCNKRNS 95

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	496	100.0	95	1 MGBB_HUMAN	O75556 homo sapien
2	281	56.7	93	1 MGBB_HUMAN	Q13296 homo sapien
3	195.5	39.4	95	1 PSC2_RAT	P02780 ratu mus sapien
4	71.5	14.4	131	1 YRS3_MCTG	O33301 mycobacteri
5	70	14.1	303	1 FTST_RICPR	O05948 rickettsia
6	67.5	13.6	112	1 PSC2_RAT	P02781 ratu mus sapien
7	67.5	13.6	268	1 TRPC_ACICA	P40911 acinetobact
8	66.5	13.4	234	1 RADC_HAEIN	P48352 haemophilus
9	65.5	13.2	162	1 IL15_MOUSE	P06420 human papil
10	65.5	13.2	603	1 VEL_HPY08	P06420 human papil
11	65	13.1	2144	1 GLT1_YEAST	Q12680 saccharomyc
12	64.5	13.0	153	1 Y136_UREPA	Q09496 ureaplasma
13	64.5	13.0	603	1 VEL_HPV21	P50759 human papil
14	64.5	13.0	611	1 DYNL_ARATH	P42697 arabidopsis
15	63.5	12.8	582	1 DORT_YEAST	Q04089 saccharomyc
16	63.5	12.8	1294	1 YAB3_SCHPO	O09716 schizosach
17	63	12.7	1022	1 SCAR_RICPR	O09449 rickettsia
18	62.5	12.6	91	1 UTER_RABIT	P02779 oryctolagus
19	62.5	12.6	113	1 MTH_CAIAT	P55321 gallinestes
20	62.5	12.6	223	1 YW24_YEAST	P40218 saccharomyc
21	62.5	12.6	663	1 YAZ1_SCHPO	P79005 schizosach
22	62	12.5	681	1 CA02_RABIT	O07067 o acyl-coen
23	62	12.5	938	1 V120_HSVT7	P52438 human herpe
24	62	12.5	1071	1 PR16_YEAST	P15938 saccharomyc
25	61.5	12.4	105	1 YR03_LISMO	O84365 listeria mo
26	61.5	12.4	570	1 HEM1_KULIA	P76698 kluyveromyc
27	61.5	12.4	734	1 MERE_THEMA	O92112 thermotoga
28	61.5	12.4	2190	1 CCAD_CHICK	O73700 gallus gall
29	61	12.3	63	1 YEWB_ECOLI	P76275 escherichia
30	61	12.3	605	1 VEL_HPV14	P36721 human papil
31	61	12.3	606	1 VEL_HPV14	P25542 human papil
32	61	12.3	732	1 KUB6_MOUSE	P27641 mus musculu
33	61	12.3	830	1 J1P2_MOUSE	O94969 mus musculu

34	61	12.3	1272	1 Y228_METHA	O60287 methanococ
35	60.5	12.2	105	1 Y531_LISTIN	O92748 listeria in
36	60.5	12.2	180	1 MUP2_MOUSE	P11589 mus musculu
37	60.5	12.2	430	1 SERC_ARATH	O96255 arabidopsis
38	60.5	12.2	643	1 VEL_HPV42	P27221 human papil
39	60.5	12.2	648	1 TALA_POVMK	P24597 mouse papil
40	60.5	12.2	684	1 TC10_YEAST	P50273 saccharomyc
41	60.5	12.2	1510	1 CCAD_MESAU	O99244 mesocricetu
42	60	12.1	351	1 PERB_ARMRU	P15223 armocacta t
43	60	12.1	604	1 VEL_HPV12	O05134 human papil
44	60	12.1	606	1 VEL_HPV05	P06920 human papil
45	60	12.1	790	1 V90K_AWVLE	P03593 alfalfa mos

ALIGNMENTS

RESULT 1
MGBB_HUMAN STANDARD: PRT: 95 AA.
AC O75556;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mammaglobin B precursor (Mammaglobin 2) (Lipophilin C) (Lactroglobin)
DE (Secretoglobin family 2A member 1).
GN SCGB2A1 OR MGB2 OR UGB3 OR LIPHC.
OS Homo sapiens (Human).
OC Eukaryota; Euteleostomi; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99167354; PubMed-9806831;
RA Becker R.M., Darrow C., Zimonjic D.B., Popescu N.C., Watson M.A.,
RA Fleming T.P.;
RT "Identification of mammaglobin B, a novel member of the uteroglobin
RT gene family.";
RL Genomics 54:70-78(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99167354; PubMed-10066439;
RA Zhao C., Nguyen T., Yustitov T., Glasgow B.J., Lehrer R.I.;
RT "Lipophilin: human peptides homologous to rat prostatein.";
RL Biochem. Biophys. Res. Commun. 256:147-155(1999).
RN [3]
RP SEQUENCE OF 19-85.
RC TISSUE-Tears:
RA MEDLINE-98163342; PubMed-9504814;
RA Molloy M.P., Bolis S., Herbert B.R., Ou K., Tyler M.I., van Dyk D.D.,
RA Wilcox M.D., Gooley A.A., Williams K.L., Morris C.A., Walsh B.J.;
RT "Establishment of the human reflex tear two-dimensional polyacrylamide
RT gel electrophoresis reference map: new proteins of potential
RT diagnostic value.";
RL Electrophoresis 18:2811-2815(1997).
RN [4]
RP SEQUENCE OF 19-46 AND 60-78, AND MASS SPECTROMETRY.
RX MEDLINE-98385871; PubMed-9720917;
RA Lehrer R.I., Xu G., Abduragimov A., Dinh N.N., Ou X.-D., Martin D.,
RA Glasgow B.J.;
RT "Lipophilin, a novel heterodimeric protein of human tears.";
RL FEBS Lett. 432:163-167(1998).
CC -1- FUNCTION: MAY BIND ANDROGENS AND OTHER STEROIDS, MAY ALSO BIND
CC ESTRADIOL. A CHEMOTHERAPEUTIC AGENT USED FOR PROSTATE CANCER.
CC MAY BE UNDER TRANSCRIPTIONAL REGULATION OF STEROID HORMONES.
CC -1- SUBUNIT: HETERODIMER OF A LIPOPHILIN A AND A LIPOPHILIN C
CC (MAMMAGLOBIN B) MONOMER ASSOCIATED HEAD TO HEAD.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN THYMUS, TRACHEA, KIDNEY, STEROID
CC RESPONSIVE TISSUES (PROSTATE, TESTIS, UTERUS, BREAST AND OVARY),
CC AND SALIVARY GLAND.
CC -1- MASS SPECTROMETRY: MW=8854.94; METHOD=Electrospray; RANGE=19-95.
CC -1- SIMILARITY: BELONGS TO THE UTEROGLOBIN FAMILY. LIPOPHILIN
CC SUBFAMILY.

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: AF071219; AAC79996.1; -
DR EMBL: AJ224173; CAA1865.1; -
DR Genew: HGNC:7051; SCGB2A2.
DR MIM: 604398; -
DR InterPro: IPR003627; Mamgb/prostatn.
DR InterPro: IPR000329; Uteroglobln.
DR Pfam: PF01099; Uteroglobln.1.
DR Prodom: PD029354; Mamgb/prostatn.1.
DR PROSITE: PS00403; UTEROGLOBIN_1; FALSE-NEG.
DR PROSITE: PS00404; UTEROGLOBIN_2; FALSE-NEG.
DR Signal: glycoprotein.
KW SIGNAL 1
FT CHAIN 19 95 MAMMAGLOBIN B.
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 95 AA; 10884 MW; 0719738289f89f8D CRC64;

Query Match 100.0%; Score 496; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 4e-44;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLWLVMLAALLHCYACDSCCKLEEDVKEKTSISPEYKELLOEFIDSDAAAEAMG 60
DB 1 MKLWLVMLAALLHCYACDSCCKLEEDVKEKTSISPEYKELLOEFIDSDAAAEAMG 60
OY 61 KRKQCFLNQSHRTKNGFLMMHTVYDSINCMKSN 95
DB 61 KRKQCFLNQSHRTKNGFLMMHTVYDSINCMKSN 95

RESULT 2
MGBA_HUMAN STANDARD; PRT; 93 AA.
ID MGBA_HUMAN 013296;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mammaglobin A Precursor (Mammaglobin 1) (Secretoglobin family 2A member 2).
GN SCGB2A2 OR MGB1 OR UGB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Breast;
RX MEDLINE=96223698; Pubmed=8631025;
RA Watson M.A., Fleming T.P.;
RT "Mammaglobin, a mammary-specific member of the uteroglobln gene family, is overexpressed in human breast cancer.";
RL Cancer Res. 56:860-865(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96147371; Pubmed=9488047;
RA Watson M.A., Darrow C., Zimonjic D.B., Popescu N.C., Fleming T.;
RT "Structure and transcriptional regulation of the human mammaglobln gene, a breast cancer associated member of the uteroglobln gene family localized to chromosome 11q13.";
RL Oncogene 16:817-824(1998).
RN [3]
RP TISSUE SPECIFICITY: MAMMARY-SPECIFIC. OVER-EXPRESSED IN BREAST CANCER.
CC -1- SIMILARITY: BELONGS TO THE UTEROGLOBIN FAMILY. LIPOPHILIN SUBFAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: U33147; AAC50375.1; -
DR EMBL: AF015224; AAC39608.1; -
DR Genew: HGNC:7050; SCGB2A2.
DR MIM: 605562; -
DR InterPro: IPR003627; Mamgb/prostatn.
DR InterPro: IPR000329; Uteroglobln.
DR Pfam: PF01099; Uteroglobln.1.
DR Prodom: PD029354; Mamgb/prostatn.1.
DR SMART: SM00086; UTRG.1.
DR PROSITE: PS00403; UTEROGLOBIN_1; FALSE-NEG.
DR PROSITE: PS00404; UTEROGLOBIN_2; I.
DR Signal: glycoprotein.
KW SIGNAL 1
FT CHAIN 19 93 MAMMAGLOBIN A.
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 93 AA; 10499 MW; 2895E8C43BF053E2 CRC64;

Query Match 56.7%; Score 281; DB 1; Length 93;
Best Local Similarity 58.7%; Pred. No. 3.8e-22;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

OY 1 MKLWLVMLAALLHCYACDSCCKLEEDVKEKTSISPEYKELLOEFIDSDAAAEAMG 60
DB 1 MKLWLVMLAALLHCYACDSCCKLEEDVKEKTSISPEYKELLOEFIDSDAAAEAMG 60
OY 61 KRKQCFLNQSHRTKNGFLMMHTVYDSINCMKSN 92
DB 61 KRKQCFLNQSHRTKNGFLMMHTVYDSINCMKSN 92

RESULT 3
PSC3_RAT STANDARD; PRT; 95 AA.
ID PSC3_RAT P02780; 063463;
AC P02780; 063463 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DE Prostatic steroid-binding protein C3 chain precursor (Prostatein peptide C3).
DE Peptide C3).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83082848; Pubmed=6294095;
RA Parker M.G., White R., Hurst H., Needham M., Tilly R.;
RT "Prostatic steroid-binding protein. Isolation and characterization of C3 genes.";
RL J. Biol. Chem. 258:12-15(1983).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=83238526; Pubmed=6190812;
RA Viskochil D.H., Perry S.T., Lea O.A., Stafford D.W., Wilson E.M., French F.S.;
RT "Isolation of two genomic sequences encoding the Mr = 14,000 subunit of rat prostatein.";
RL J. Biol. Chem. 258:8861-8866(1983).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=92165796; Pubmed=1537831;
RA Tan J.A., Marschke K.B., Ho K.-C., Perry S.T., Wilson E.M., French F.S.;
RT "Response elements of the androgen-regulated C3 gene.";

```

```

RL J. Biol. Chem. 267:4456-4466(1992).
RN [4]
RP ERRATUM.
RX MEDLINE-92218467; PubMed-1339454;
RA Tan J.A., Marschke K.B., Ho K.C., Perry S.T., Wilson E.M.,
RA French F.S.;
RL J. Biol. Chem. 267:7958-7958(1992).
RN [5]
RP SEQUENCE OF 19-95.
RX MEDLINE-8186769; PubMed-7014218;
RA Peeters B., Rombauts W., Mous J., Heyns W.;
RT *Structural studies on rat prostatic binding protein. The primary
RT structure of its glycosylated component C3.*;
RL Eur. J. Biochem. 115:115-121(1981).
CC -1- FUNCTION: PART OF PROSTATEIN WHICH IS THE MAJOR SECRETORY
CC GLYCOPROTEIN OF VENTRAL PROSTATE GLAND. STEROID-BINDING PROTEIN;
CC CAN BIND NON-POLAR STEROIDS. CHOLESTEROL AND A GROUP OF SMALL
CC PROLINE-RICH PEPTIDES.
CC -1- SUBUNIT: PROSTATEIN IS COMPOSED OF THREE DIFFERENT PEPTIDES CALLED
CC C1, C2 AND C3. THESE FORM COVALENT C1:C3 (P) AND C2:C3 (S)
CC HETERODIMERS WHOSE NONCOVALENT ASSOCIATION FORMS TETRAMERIC
CC (C1:C3/C3:C2) PROSTATEIN MOLECULES.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: VENTRAL PROSTATE.
CC -1- INDUCTION: ANDROGEN DEPENDENT. AS SHOWN BY THE DECREASE IN THE
CC LEVEL OF THE PROTEIN FOLLOWING CASTRATION.
CC -1- MISCELLANEOUS: C3 IS ENCODED BY TWO DIFFERENT GENES.
CC -1- SIMILARITY: BELONGS TO THE UTEROGLOBIN FAMILY. LIPOPHILIN
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V01263; CAA24577.1; -
DR EMBL: M71245; AAA41965.1; -
DR PIR: A03250; BORT3.
DR InterPro: IPR003627; Mambo/prostratn.
DR InterPro: IPR000329; uteroglobln.
DR Pfam: PF01099; uteroglobln.1.
DR PRODOM: PD028354; Mamgd/prostatn; 1.
DR PROSITE: PS00403; UTEROGLOBIN_1; FALSE_NEG.
DR PROSITE: PS00404; UTEROGLOBIN_2; 1.
DR Signal: Glycoprotein; Steroid-binding.
FT SIGNAL 1 18
FT CHAIN 19 95
FT PROSTATIC STEROID-BINDING PROTEIN C3
FT CARBOHD 35 35
FT CONFLICT 53 53
FT CONFLICT 79 79
FT CONFLICT 79 79
FT SEQUENCE 95 AA: 10730 MW; F7F7F1A0C882E375 CRC64;
SQ
Query Match 39.4%; Score 195.5; DB 1; Length 95;
Best Local Similarity 35.8%; Pred. No. 2,2e-13;
Matches 34; Conservative 30; Mismatches 30; Indels 1; Gaps 1;
QY 1 MRLMVLMLALLLHCTAD-SGCKLEDVETKINSIDISPEYKELLOFTDSDAAPAM 59
DB 1 MRLVFLFLVLPITCCVAGSGCSILDEVIRGTINSTVLHDYMKLVPRYVDHATERAV 60
QY 60 GKFKCFLNOSHRTLNKNGMLMHTVYDSICNKRK 94
DB 61 KQFKCFLDQDTKLTENVGVMEAFENSECQOPS 95

```

```

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV2759c.
GN RV2759C OR MT2829 OR MTV002.24C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-H37RV;
RX MEDLINE-98295987; PubMed-9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Horsley T., Jagels K., Krogh A., McLaren J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Ruster J.E., Taylor K., Skellon S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-CDC 1551 / Oshkosh;
RX Flatschmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RX Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RX Kolony J.F., Nelson W.C., Umayam L.A., Emlaeva M.D., Salzberg S.L.,
RX Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikala A.,
RA Bichel W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE UPF0110 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL008967; CAA15555.1; -
DR EMBL: AE007110; AAK47148.1; -
DR TIGR: MT2829; -
DR Tuberculist: RV2759C; -
DR InterPro: IPR005342; PF03655.
DR Pfam: PF03655; UPF0110; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 131 AA: 14372 MW; 9C12AE0AC674065B CRC64;
Query Match 14.4%; Score 71.5; DB 1; Length 131;
Best Local Similarity 21.2%; Pred. No. 1.5;
Matches 21; Conservative 18; Mismatches 35; Indels 25; Gaps 1;
QY 5 MYLMLALLLHCTADSGCKLEDVETKINSIDISPEYKELLO----- 47
DB 1 MRLVFLFLVLPITCCVAGSGCSILDEVIRGTINSTVLHDYMKLVPRYVDHATERAV 60
QY 48 -----EFIDSDAAAEAMGKFKCFLNOSHRTLNKNG 78
DB 61 LDDVCIQVEAVDAQARVAQAQAYRDYGRGSGHAPRLNLG 99

```

```

RESULT 4
YR59_MYCTU STANDARD: PRT; 131 AA.
ID YR59_MYCTU
AC 033301.

```

```

RESULT 5
FTSY_RICPR STANDARD: PRT; 303 AA.
ID FTSY_RICPR
AC 005948;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)

```

DT 16-OCT-2001 (Rel.. 40, last annotation update)
 DE Cell division protein first homolog.
 GN P157 OR RP775.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 CC Rickettsiaceae; Rickettsiidae; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE=9741951; PubMed=9274032;
 RA Andersson J.O., Andersson S.G.E.;
 RT "Genomic rearrangements during evolution of the obligate
 intracellular parasite Rickettsia prowazekii as inferred from an
 analysis of 52015 bp nucleotide sequence.";
 RL Microbiology 143:2783-2795(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE=9903499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 R. Sichelitz-Ponten T., Almarik U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 mitochondria.";
 RL Nature 396:133-140(1998).
 RN [3]
 RP SEQUENCE OF 168-303 FROM N.A.
 RC STRAIN=Madrid E, and B;
 RX MEDLINE=9941641; PubMed=10486973;
 RA Andersson J.O., Andersson S.G.E.;
 RT "Genome degradation is an ongoing process in Rickettsia.";
 RL Mol. Biol. Evol. 16:1178-1191(1999).
 CC -1- FUNCTION: FUNCTIONAL HOMOLOG OF SRP RECEPTOR, PROBABLY INVOLVED IN
 THE RECEPTION AND INSERTION OF A SUBSET OF PROTEINS AT THE
 MEMBRANE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y11784; CAA72477.1; -;
 CC EMBL: AJ235273; CAA15202.1; -;
 CC EMBL: AJ238755; CAB56084.1; -;
 CC EMBL: AJ238756; CAB56088.1; -;
 CC HSSP: P10121; IPTS.
 CC InterPro: IPR004390; FTSY.
 CC InterPro: IPR000897; SRP54.
 CC InterPro: IPR004781; SRP_dock.
 CC Pfam: PF00448; SRP54; 1.
 CC Pfam: PF02881; SRP54_N; 1.
 CC ProDom: PD000819; SRP54; 1.
 CC TIGRfam: TIGR00064; ftsy; 1.
 CC TIGRfam: TIGR00961; 3a0501503; 1.
 CC PROSITE: PS00300; SRP54; 1.
 CC Signal recognition particle; GTP-binding; RNA-binding; Membrane;
 CC Cell division; Complete proteome.
 CC NP_BIND 108 115 GTP (BY SIMILARITY).
 CC NP_BIND 190 194 GTP (BY SIMILARITY).
 CC NP_BIND 254 257 GTP (BY SIMILARITY).
 CC SEQUENCE 303 AA; 33313 MW; BF3F7EB9383B7007 CAC64;
 Query Match 14.18; Score 70; DB 1; Length 303;
 Best Local Similarity 35.28; Pred. No. 5.1;
 Matches 19; Conservative 10; Mismatches 17; Indels 8; Gaps 1;

Db 37 LNELELLISSDISISVYTHIEFKVNFKPTIDSDVKEAIAIKLIDQLSKS 90
 RESULT 6
 AC PSC2_RAT STANDARD; PRT; 112 AA.
 ID PSC2_RAT
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Prostatic steroid-binding protein C2 chain precursor (Prostatein
 DE peptide C2).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8714648; PubMed=2881277;
 RA Delaey B., Dirckx L., Decourt J.-L., Claessens F., Peeters B.,
 RA Rombaux W.;
 RT "Rat prostatic binding protein: the complete sequence of the C2 gene
 RT and its flanking regions.";
 RL Nucleic Acids Res. 15:1627-1641(1987).
 RN [2]
 RP SEQUENCE OF 21-112.
 RX MEDLINE=83209619; PubMed=6343081;
 RA Peeters B., Heyns W., Mous J., Rombaux W.;
 RT "Structural studies on rat prostatic binding protein. The primary
 RT structure of component C2 from subunit S.";
 RL Eur. J. Biochem. 132:669-679(1983).
 RN [3]
 RP SEQUENCE OF 1-100 FROM N.A.
 RX MEDLINE=82220075; PubMed=6896362;
 RA Parker M., Needham M., White R.;
 RT "Prostatic steroid binding protein: gene duplication and steroid
 RT binding.";
 RL Nature 298:92-94(1982).
 CC -1- FUNCTION: PART OF PROSTATEIN WHICH IS THE MAJOR SECRETORY
 CC GLYCOPROTEIN OF VENTRAL PROSTATE GLAND.
 CC -1- SUBUNIT: PROSTATEIN IS COMPOSED OF THREE DIFFERENT PEPTIDES CALLED
 CC C1, C2 AND C3. THESE FORM COVALENT C1:C3 (P) AND C2:C3 (S)
 CC HETERODIMERS WHOSE NONCOVALENT ASSOCIATION FORMS TETRAMERIC
 CC (C1:C3/C3:C2) PROSTATEIN MOLECULES.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PTM: LINKED BY THREE DISULFIDE BONDS TO C3.
 CC -1- MISCELLANEOUS: THE HETERODIMER CAN BIND NON-POLAR STEROIDS,
 CC CHOLESTEROL AND A GROUP OF SMALL PROLINE-RICH PEPTIDES.
 CC -1- SIMILARITY: BELONGS TO THE UTEROGLOBIN FAMILY. LIPOPHILIN
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X03034; CAA28708.1; -;
 CC EMBL: V01256; CAA24569.1; -;
 CC EMBL: J00776; AAA51641.1; -;
 CC PIR: A03251; BORT2.
 CC PIR: A26671; A26671.
 CC InterPro: IPR000329; uteroglobin.
 CC PROSITE: PS00403; UTEROGLOBIN_1; FALSE_NEG.
 CC PROSITE: PS00404; UTEROGLOBIN_2; FALSE_NEG.
 CC Steroid-binding; Signal.
 CC SIGNAL 1 20
 CC CHAIN 21 112
 CC MOD_RES 21 21 PROSTATIC STEROID-BINDING PROTEIN C2
 CC CHAIN.
 CC BLOCKED.

FT DISULFID 28 INTERCHAIN (WITH C3) (PROBABLE).
 FT DISULFID 69 INTERCHAIN (WITH C3) (PROBABLE).
 FT DISULFID 92 INTERCHAIN (WITH C3) (PROBABLE).
 FT CONFLICT 26 MISSING (IN REF. 3).
 FT CONFLICT 88 I -> T (IN REF. 3).
 FT CONFLICT 96 VMLQINFGKWFSEIN -> YGK (IN REF. 3).
 SQ SEQUENCE 112 AA; 112828 MW; DA65A6A82E677864 CRC64;

Query Match 13.6%; Score 67.5; DB 1; Length 112;
 Best Local Similarity 27.0%; Pred. No. 3; 2;
 Matches 27; Conservative 20; Mismatches 36; Indels 17; Gaps 5;

OY 4 LMLVLAALALHCY2DSG-----CKLEDM-VKRTINSDISIPEYKELQEFIDSDAA 55
 DB 3 LSLCLTLVVCYFANGQTFAGVCQALQDVTITFLNPE--BELKRELEFADAPPEA 59
 OY 56 AEAMGKFKGCLNGHRTLKAPGLMMHT--VYDSITCMNK 93
 DB 60 VEANLKYKRCI---NKIMYGDRLSMGTSLVFTMLKDVK 95

RESULT 7
 TRPC_ACICA STANDARD; PRT: 268 AA.

ID TRPC_ACICA STANDARD; PRT: 268 AA.
 AC P00911:
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Indole-3-glycerol phosphate synthase (EC 4.1.1.48) (IGPS).
 GN TRPC.
 OS Acinetobacter chloaceticus.
 OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 CC Acinetobacter.
 CC NCBI_Taxid=471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88174326; PubMed=6599977;
 RA Kaplan J.B., Goncharoff P., Seibold A.M., Nichols B.P.;
 RT "Nucleotide sequence of the Acinetobacter calcoaceticus trpDC gene cluster.";
 RT Mol. Biol. Evol. 1:456-472(1984).

CC -1- CATALYTIC ACTIVITY: 1-(2-carboxyphenylamino)-1-deoxy-D-ribose 5-phosphate - 1 (indol-3-yl)glycerol 3-phosphate + CO(2) + H(2)O.
 CC -1- PATHWAY: tryptophan biosynthesis, fourth step.
 CC -1- SIMILARITY: BELONGS TO THE TRPC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; M36636; AAA21905.1; -;
 DR PIR; A01133; GMEKCC.
 DR HSSP; P00909; 1P11.
 DR InterPro: IPR001468; IGPS.
 DR Pfam: PF00218; IGPS; 1;
 DR ProDom: PD001511; IGPS; 1.
 DR ProSite: PS00614; IGPS; 1.
 KW Tryptophan biosynthesis; Lyase; Decarboxylase.
 SQ SEQUENCE 268 AA; 30216 MW; 8DBAC0505D5A7527 CRC64;

Query Match 13.6%; Score 67.5; DB 1; Length 268;
 Best Local Similarity 26.4%; Pred. No. 8;
 Matches 19; Conservative 16; Mismatches 28; Indels 9; Gaps 2;

OY 12 LLLHLYADSGCKLLEDVKTINSDISIPEYKELQEFIDSDAAAEAMGKFKGCLNGSH 71
 DB 143 LTVACLSD---QOLEMSKTAFFEYDL-----DVLVEVHDEQLERALKLSEDCILGVNN 193
 OY 72 RFLKNFGIMMHT 83

Db 194 RLKTFEYDLNLT 205

RESULT 8

ID RADC_HAEIN STANDARD; PRT: 234 AA.
 AC P44952:

DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA repair protein radC homolog.
 GN RADC OR H10952.
 OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.

CC NCBI_Taxid=727;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shirley R., Liu L.T., Glodok A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Ufferback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C., RAE Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.";
 RT Science 269:496-512(1995).

CC -1- FUNCTION: INVOLVED IN DNA REPAIR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RADC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U32776; AAC22613.1; -;
 DR TIGR; H10952; -;

DR InterPro: IPR001405; RADC.
 DR ProDom: PD007415; RADC; 1.
 DR TIGRFAMS: TIGR00608; RADC; 1.
 DR ProSite: PS01302; RADC; 1.
 KW DNA repair; Complete proteome.
 SQ SEQUENCE 234 AA; 26774 MW; 25472EB289E5DA69 CRC64;

Query Match 13.4%; Score 66.5; DB 1; Length 234;
 Best Local Similarity 26.6%; Pred. No. 8;
 Matches 21; Conservative 15; Mismatches 26; Indels 17; Gaps 3;

OY 24 LLEDVKTINSDISIPEYKELQEFIDSDAAAEAMGKFKGCLNGSHRTLKNGIMMHT 83
 DB 103 LKQDMLSTPILND--PEYKVL--FLTLQLQHEREVPMVLELNDQHLIKERLPIGT 156

OY 84 VY-----DSIMCN 91
 DB 157 IYSAVYPREITKEALYCN 175

RESULT 9

ID ILIS_MOUSE STANDARD; PRT: 162 AA.
 AC P48346;

DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)

```

DE Interleukin-15 precursor (IL-15).
CN IL15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC/REF X C57BL/6J; TISSUE=Bone marrow;
RA MEDLINE=95278940; PubMed=7759105;
RA Anderson D.M., Johnson L., Giaccum M.B., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A., Valentine V., Kirstein M.N., Shapiro D.N., Morris S.W.,
RA Grabstein K., Cosman D.;
RA "Chromosomal assignment and genomic structure of IL15.";
RL Genomics 25:701-706(1995).
CC -1- FUNCTION: CYTOKINE THAT STIMULATES THE PROLIFERATION OF T-
CC LYMPHOCYTES. STIMULATION BY IL-15 REQUIRES INTERACTION OF IL-15
CC WITH COMPONENTS OF IL-2R. INCLUDING IL-2R BETA AND PROBABLY IL-2R
CC GAMMA BUT NOT IL-2R ALPHA.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE IL-15 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/Announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U14332; AAA75377.1; .
DR MGD: MGI:103014; IL15.
DR InterPro: IPR003443; Interleukin_15.
DR Pfam: PF02372; IL15: 1.
DR Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 29
FT PROPEP 30 48
FT CHAIN 49 162
FT DISULFID 83 133
FT DISULFID 90 136
FT DISULFID 104 104
FT CARBOHYD 108 108
FT CARBOHYD 119 119
FT SEQUENCE 162 AA; 686371498CEBB296 CRC64;
SQ
Query Match 13.28; Score 65.5; DB 1; Length 162;
Best Local Similarity 34.0%; Pred. No. 7.6;
Matches 18; Conservative 6; Mismatches 10; Indels 19; Gaps 2;
OY 18 ADSCGKLEDDWEKTIINSIPYKELLEPFIDSDAAEAAGKFGCCFLNOS 70
DB 129 AESGCKEECELEKTFTEFLDSFT-----RIVQMFINTS 162

```

```

RL J. Virol. 58:626-634(1986).
CC -1- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CC CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/Announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M12737; -; NOT_ANNOTATED_CDS.
DR PIR: A03656; MWLE8.
DR InterPro: IPR001177; Papillom_E1.
DR Pfam: PF00519; E1: 1.
DR Pfam: PF00524; E1_N: 1.
DR Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
KW Nuclear protein.
FT NP_BIND 431 438
FT SEQUENCE 603 AA; 68821 MW; 081386009DD48AD CRC64;
SQ
Query Match 13.28; Score 65.5; DB 1; Length 603;
Best Local Similarity 22.28; Pred. No. 30;
Matches 22; Conservative 21; Mismatches 23; Indels 33; Gaps 4;
OY 19 DSGCKLEDDWEKTIINSIP-----EYKLEQFIDSDAAEAAGK 61
DB 112 DSGVELPLTNNEDVSHVEVPALDSRPEDEGSGALDIDYALLR---SSYTKATLMAK 168
OY 62 FKQCKFLN-----QSHRLKNFGLM---MHTYYDS 87
DB 169 FKFAFGDFNELTRQFSKYTCQNVVAAYAVADYVES 207

```

```

RESULT 11
ID GLUT1 YEAST STANDARD; PRT; 2144 AA.
AC 012680; 012290.
DT 01-NOV-1987 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Glutamate synthase [NADPH] precursor (EC 1.4.1.13) (NADPH-GOGAT).
GN GLUT1 OR YDL171C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=CN36;
RA MEDLINE=97082505; PubMed=8923741;
RA Filetici P., Martegani M.P., Valenzuela L., Gonzalez A., Ballarín P.;
RA "Sequence of the GLUT1 gene from Saccharomyces cerevisiae reveals the
RA domain structure of yeast glutamate synthase.";
RL Yeast 12:1359-1366(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RA Pohl T.M.;
RA Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: 2 L-glutamate + NAD(+) -> L-glutamate + 2-
CC oxoglutarate + NADPH.
CC -1- COFACTOR: BINDS A 3FE-4S CLUSTER; FAD AND FMN.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SIMILARITY: TO OTHER GLUTAMATE SYNTHASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC -----
DR EMBL: AEO02115; AAF30562.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 153 AA; 18775 MW; 894C11DAFF37B7AC CRC64;
OY
Query Match 13.0%; Score 64.5; DB 1; Length 153;
Best Local Similarity 30.2%; Pred. No. 9;
Matches 19; Conservative 13; Mismatches 22; Indels 9; Gaps 3;
OY 22 CKLEDMVEKTIKNSD--ISIPYKLLQEFIDSDAAAEAMGKFRQCFGLNSHRT-----L 74
DB 49 CKYVPRKISKHLNNDHLINVDYVLTREKFIKSKYEKV--KHYLRFISTENKTKOMPEL 106
OY 75 KNP 77
DB 107 INF 109
RESULT 13
VEI_HPV21 STANDARD: PRT; 603 AA.
AC PS0759;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Replication protein E1.
GN E1.
OS Human papillomavirus type 21.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=31548;
RN [1]
RA Delius H.;
RA Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
CC -I- VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CC CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@sib-sib.ch).
CC CC -----
DR EMBL: U31779; AAA79396.1; -.
DR InterPro: IPR001177: Papillom_E1.
DR Pfam: PF00519; E1.1.
DR Pfam: PF00534; E1.N.1.
KW Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
KW Nuclear protein.
KW NP_BIND 431 438 ATP (POTENTIAL).
SO SEQUENCE 603 AA; 68983 MW; 93006494BEF10339 CRC64;
OY
Query Match 13.0%; Score 64.5; DB 1; Length 603;
Best Local Similarity 26.3%; Pred. No. 38;
Matches 26; Conservative 15; Mismatches 23; Indels 35; Gaps 6;
OY 19 DSG--CKLEDMVEKTIKNSDISIP-----EYKELLQEFIDSDAAEAM 59
DB 112 DSGECLTIN---EEVSSSEVPEALDSQPAEAOLGVDIHNYKELLR---SNKRAILM 165
OY 60 GKFRQCF---LNSHRTLKNFQ---LMMHYVYDSI 88
DB 166 AKREFFGVGFNDLTROFKSYKTCCKMANVLSYAVHDL 204
RESULT 14
OYNL_ARATH

```

```

ID DYNL_ARATH STANDARD: PRT: 611 AA.
AC P42697;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DYNALIN-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucotids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RA Hwang I., Goodman H.M., Yoon H.W., Yoon J.-H., Cho M.J.;
RT "Cloning and characterization of a cDNA clone encoding dynamin-like
RT GTP binding protein in Arabidopsis thaliana."
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -! SIMILARITY: BELONGS TO THE DYNAMIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: L36939; AAB63528.1; -.
DR InterPro: IPR001401; Dynamin.
DR InterPro: IPR000375; Dynamin_central.
DR InterPro: IPR003130; GED.
DR Pfam: PF00350; dynamin.1.
DR Pfam: PF01031; dynamin_2.1.
DR Pfam: PF02212; GED.1.
DR PRINTS: PR00195; DYNAMIN.
DR SMART: SM00053; DYNM.1.
DR SMART: SM00302; GED.1.
DR PROSITE: PS00410; DYNAMIN.1.
DR Motor Protein: GTP-binding.
KW NP_BIND 41 48 GTP (POTENTIAL).
FT NP_BIND 142 146 GTP (POTENTIAL).
FT NP_BIND 211 214 GTP (POTENTIAL).
FT FT 611 AA: 68509 MW: EDE98B79B1BC3D5 CRC64.
SQ
SEQUENCE 611 AA: 68509 MW: EDE98B79B1BC3D5 CRC64.

Query Match 13.08; Score 64.5; DB 1; Length 611;
Best Local Similarity 20.7%; Pred. No. 38;
Matches 17; Conservative 21; Mismatches 23; Indels 21; Gaps 3;

OY 24 LLEDVETKINSISIPYKELDGFIDSDAAEAMGKPK-----OC----- 65
DB 432 ILKDLVHSVMEYELKQPALREYV--TNAALIESLDKMRGSKATLQLVMECSYLTV 489
OY 66 -FLNOSHRTLNFGIMHTTYVD 86
DB 490 DFRRLPDQVEKGNPTHSIFD 511

RESULT 15
DOT1_YEAST STANDARD: PRT: 582 AA.
AC Q04089;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Disrupter of telomere silencing protein 1.
DE DONT1 OR PCH1 OR YDR440W OR D9461.26.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A. AND FUNCTION.

```

```

RX MEDLINE=96429517; PubMed=9755194;
RA Singer M.S., Kahana A., Wolf A.J., Meisinger L.L., Peterson S.E.,
RA Goggin C., Mahowald M., Gottschling D.E.;
RT "Identification of high-copy disruptors of telomeric silencing in
RT Saccharomyces cerevisiae."
RL Genetics 150:613-632(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
RA Carpenter J., Chen E., Cherry J.W., Chung E., Duncan M.,
RA Hunkeler-Smith S., Hyman R., Komp C., Laskari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Oetner P., Oh C., Petel F.X.,
RA Roberts D., Schramm S., Schroeder M., Shogen T., Shroff N.,
RA Winant A., Yellon M., Botstein D., Davis R.W.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: Has a role in telomere silencing, which is the
CC repression of chromatin structure which leads to a stop in the
CC transcription of nearby genes.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U33007; AAB64868.1; -.
DR SGD: S0002848; DONT1.
DR Transcription regulation.
SQ SEQUENCE 582 AA: 66201 MW: 05CA6A68F8CBAB9A CRC64.

Query Match 12.88; Score 63.5; DB 1; Length 582;
Best Local Similarity 25.7%; Pred. No. 46;
Matches 19; Conservative 13; Mismatches 15; Indels 27; Gaps 3;

OY 10 AALLHGYADSGCKLLEDVETKINSISIPYKE-----LQGFID 51
DB 409 AALECGCALSFGECEIMDA-----SDLTLLQYEELKRCRKLGMRLNNVERSLKKSFD 462
OY 52 SDAAEAMGKPKOC 65
DB 463 NNKRVALL---PQC 473

Search completed: January 25, 2003, 13:31:47
Job time : 85 secs

```



```

O9GK61
ID O9GK61 PRELIMINARY: PRT: 93 AA.
AC O9GK61:
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Lipophilin CS.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SUBMAXILLARY;
RA Zhao C., Nguyen T.X., Lehrer R.I.;
RT "Rabbit Lipophilins."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF308620; AAG42808.1;
DR InterPro: IPR003627; Mambp/prostatn.
DR Pfam: PF01099; Uteroglobln.
DR PRINTS: PR00486; UTEROGLOBIN.
DR ProDom: PD029354; Mambp/prostatn.
DR SMART: SM00096; UTC:1.
SQ SEQUENCE 93 AA: 10609 MW: E2B015AD31B2249 CRC64:

Query Match 46.8%; Score 232; DB 6; Length 93;
Best Local Similarity 50.0%; Pred. No. 2,1e-17;
Matches 45; Conservative 20; Mismatches 25; Indels 0; Gaps 0;

OY 1 MKLWVLMALALLHCYVADSCKLLEDMVEKTLNSDISIPEKRELOEFLDSDAAAEAMG 60
DB 1 MKVIVLMALALPLYCYAGSCCPVEKVKVTKTLNSVSAEYIDLVKNIINDEPRLAVV 60
OY 61 KFKOCFLNQSHTLKNFGMLMHTVYDSIMC 90
DB 61 EKKCFLLSQSEETLNVEMETITNSKLC 90

RESULT 3
O9GK64
ID O9GK64 PRELIMINARY: PRT: 93 AA.
AC O9GK64:
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Lipophilin CL.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LACRIMAL GLAND;
RA Zhao C., Nguyen T.X., Lehrer R.I.;
RT "Rabbit Lipophilins."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF308617; AAG42805.1;
DR InterPro: IPR003627; Mambp/prostatn.
DR InterPro: IPR000329; Uteroglobln.
DR Pfam: PF01099; Uteroglobln.
DR ProDom: PD029354; Mambp/prostatn.
SQ SEQUENCE 93 AA: 10350 MW: 9F9206C44372804D CRC64:

Query Match 46.2%; Score 229; DB 6; Length 93;
Best Local Similarity 45.6%; Pred. No. 4,4e-17;
Matches 41; Conservative 26; Mismatches 23; Indels 0; Gaps 0;

OY 1 MKLWVLMALALLHCYVADSCKLLEDMVEKTLNSDISIPEKRELOEFLDSDAAAEAMG 60
DB 1 MKVIVLMALALPLYCYAGSCCVLESVLDKTDIPSVSEVDYTYLQKIITLDAKVALE 60
OY 61 KFKOCFLNQSHTLKNFGMLMHTVYDSIMC 90

```

```

DB 61 ELKOCFLSQSNETLANVKVLEAVFDSLVC 90

RESULT 4
O9GK62
ID O9GK62 PRELIMINARY: PRT: 93 AA.
AC O9GK62:
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Lipophilin CP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PROSTATE;
RA Zhao C., Nguyen T.X., Lehrer R.I.;
RT "Rabbit Lipophilins."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF308619; AAG42807.1;
DR InterPro: IPR003627; Mambp/prostatn.
DR InterPro: IPR000329; Uteroglobln.
DR Pfam: PF01099; Uteroglobln.
DR ProDom: PD029354; Mambp/prostatn.
SQ SEQUENCE 93 AA: 10352 MW: C9DC35B17D372F32 CRC64:

Query Match 45.8%; Score 227; DB 6; Length 93;
Best Local Similarity 50.0%; Pred. No. 7,2e-17;
Matches 45; Conservative 16; Mismatches 29; Indels 0; Gaps 0;

OY 1 MKLWVLMALALLHCYVADSCKLLEDMVEKTLNSDISIPEKRELOEFLDSDAAAEAMG 60
DB 1 MKCVTALMALALPLYCTAGSCQQLDDMYTKTLDQSISLDYHNPFKNLSSGAAAEAMVK 60
OY 61 KFKOCFLNQSHTLKNFGMLMHTVYDSIMC 90
DB 61 DEKOCFLMQSNETLNNIKVFLETVYNSPFC 90

RESULT 5
O9JHB9
ID O9JHB9 PRELIMINARY: PRT: 95 AA.
AC O9JHB9:
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Prostatic steroid binding protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=84057754; PubMed=6685625;
RA Hurst H.G., Parker M.G.;
RT "Rat prostatic steroid binding protein: DNA sequence and transcript
maps of the two C3 genes."
RL EMBL J. 2:769-774(1983).
DR EMBL: V01260; CAB75892.1;
DR EMBL: V01261; CAB75892.1; JOINED.
DR EMBL: V01262; CAB75892.1; JOINED.
DR InterPro: IPR003627; Mambp/prostatn.
DR InterPro: IPR000329; Uteroglobln.
DR Pfam: PF01099; Uteroglobln.
DR ProDom: PD029354; Mambp/prostatn.
SQ SEQUENCE 95 AA: 10622 MW: R209F1BE177C52N4 CRC64:

Query Match 38.6%; Score 191.5; DB 11; Length 95;
Best Local Similarity 34.7%; Pred. No. 4,7e-13;
Matches 33; Conservative 31; Mismatches 30; Indels 1; Gaps 1;

```

[illegible][illegible]

```

OC Eukaryota: Fungi: Ascomycota: Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes;
OX NCBI_TaxID=4896;
RM 11
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Murphy L., Harris D., Barrell B.G., Rajandream M.A., Wood V.;
RL Submitted (SEP-1997) to the EMBL/Genbank/DBPST databases.
CC -1 SIMILARITY: BELONGS TO THE UPF0029 FAMILY.
DR EMBL; 26878; CAB1676.1;
DR InterPro: IPR001498; UPF0029.
DR Pfam: PF01205; UPF0029; 1.
KM Hypothetical protein.
FT DOMAIN 276 280
SQ SEQUENCE 280 AA; 31485 MW; 90PDE06A14091219 CRC64;

Query Match 15.48; Score 76.5; DB 3; Length 280;
Best Local Similarity 30.38; Pred. No. 3.2;
Matches 23; Conservative 10; Mismatches 28; Indels 15; Gaps 3;

OY 24 LLEDMVERTINSIDISPEYKELQEFIDSD--AAAEAMGKFKQ-----CFLNQSHT 73
DB 68 LAEDVLVSATGDCYCFISYMDLLKELVDIDAEOQAARESKLQESDKETPMLNKSHYV 127
OY 74 LKNFGGLMHHYYSIM 69
DB 128 AKT-----PEIDEPW 138

RESULT 10
OY 09A622 PRELIMINARY; PRT; 212 AA.
AC 09A622:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE D-alanyl-D-alanine dipeptidase.
CN CC2273.
OS Caulobacter crescentus.
OC Bacteria: Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;
RM 11
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Pirochka J., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Kolonay J.E., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterbeck T., Tran K., Wolf A., Vamathavan J., Ermolaeva M., White S.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005897; AA02444.1;
DR TIGR; CC2273;
DR InterPro: IPR000755; Peptidase_M45.
DR Pfam: PF01427; Peptidase_M45; 1.
DR ProDom: PD013478; Peptidase_M45; 1.
KM Complete proteome.
SQ SEQUENCE 212 AA; 23874 MW; 82846D0F44F07028 CRC64;

Query Match 15.34; Score 76; DB 16; Length 212;
Best Local Similarity 29.94; Pred. No. 2.7;
Matches 26; Conservative 13; Mismatches 16; Indels 32; Gaps 5;

OY 18 ADSCGLLEDMVEKT-----INSDISPEYK---LLQEFIDSDAAAEAMG 60
DB 19 AETGKLAISLDVDTTTPGKIDIRYAGANNFMGIPLEYKSAVLQR-----PAAEALG 73
OY 61 KFKQCFLNQSHRTL--KNFGMLMHYTV 85

```

```

DB 74 RI-----HRAUAGYGLLIDHAY 92

RESULT 11
OY 09CAN6 PRELIMINARY; PRT; 590 AA.
AC 09CAN6:
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 66.1 kDa protein.
GN F16M19.15.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RM 11
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.N.,
RA Sun H., Taiton L.J., Tambunga G., Tortum M.J., Town C.D.,
RA Uterbeck T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana."
RL Nature 408:816-820(2000).
DR EMBL; AC010795; AGS1617.1;
DR InterPro: IPR002885; PPR.
DR Pfam: PF01535; PPR; 15.
DR TRFAMS; TIGR00756; PPR; 13.
KM Hypothetical protein.
SQ SEQUENCE 590 AA; 66071 MW; 596067345518244F CRC64;

Query Match 14.84; Score 73.5; DB 10; Length 590;
Best Local Similarity 32.38; Pred. No. 16;
Matches 21; Conservative 8; Mismatches 23; Indels 13; Gaps 3;

OY 13 LHCYADSG-----KLEDMVEKTINSIDISPEYKELQEFIDSDAAAE--NKG 61
DB 291 LTSCANCNGRSDASRLSDMLKXINPDIVF--FNALIDAFYKEGLVEKLYDEYWK 348
OY 62 FKQCF 66
DB 349 SKHCF 353

RESULT 12
OY 057483 PRELIMINARY; PRT; 1688 AA.
AC 057483:
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Voltage-dependent L-type calcium channel, ALPHA-1S subunit
DE (FALPHA1S).
OS Rana catesbeiana (Bull frog).
OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.

```


[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2003, 13:18:14 : Search time 139 Seconds
(without alignments)
91.071 Million cell updates/sec

Title: US-09-806-302a-2
Perfect score: 496
Sequence: 1 MLLMLVLAALLHCYADS.....NFGIMHTYDSIMCNKSN 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

A.GeneSeq_101002:*

1: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	496	100.0	95	18 AAW35804	Human endometrial
2	496	100.0	95	20 AAY02590	A human mammaryglobl
3	496	100.0	95	21 AAB03769	Human endometrial
4	496	100.0	95	21 AAY92226	Human endometrial
5	496	100.0	95	21 AAY92237	Mammaryglobln homolo
6	496	100.0	95	21 AAY65394	Human 5' EST relat
7	496	100.0	95	22 AAB31682	An endometrial spe
8	496	100.0	95	23 AAO20555	Protein of human l
9	496	100.0	95	23 ABB09635	Human endometrial
10	496	100.0	108	20 AAY60038	Human endometrium

11	390	78.6	74	19 AAW61649	Non-ocular disease
12	283	57.1	93	23 AA083620	Human PRO protein,
13	281	56.7	93	18 AAW10179	Mammary-specific s
14	281	56.7	93	19 AAW59777	Amino acid sequenc
15	281	56.7	93	19 AAW48432	Mammaryglobln protei
16	281	56.7	93	20 AAY01718	Mammaryglobln, a mam
17	281	56.7	93	21 AAB13786	Human mammaryglobln
18	281	56.7	93	21 AAY84622	Amino acid sequenc
19	281	56.7	93	22 AAB07517	Human mammaryglobln
20	281	56.7	93	22 AAB51127	Human mammaryglobln
21	281	56.7	410	22 AA033358	Human breast cance
22	281	56.7	743	22 AA033358	Human breast cance
23	281	56.7	1095	22 AA033357	Human breast cance
24	280	56.5	93	22 AAE07531	Human mammaryglobln
25	279	56.2	93	22 AAE07532	Human mammaryglobln
26	277	55.8	93	22 AAE07534	Human mammaryglobln
27	277	55.8	93	22 AAE07535	Human mammaryglobln
28	277	55.8	93	22 AAE07536	Human mammaryglobln
29	273	55.0	93	22 AAE07530	Human mammaryglobln
30	272	54.8	93	22 AAE07532	Human mammaryglobln
31	269	54.2	93	22 AAE07533	Human mammaryglobln
32	265.5	53.5	90	22 AAE07528	Human mammaryglobln
33	265.5	53.5	90	22 AAE07537	Human mammaryglobln
34	206	41.5	74	21 AAY84624	Amino acid sequenc
35	200	40.3	220	22 AAO22141	Rai2-mammaryglobln f
36	95	19.2	33	19 AAW48433	Mammaryglobln synthe
37	77	15.5	21	22 AAB51114	Mammaryglobln monocl
38	77	15.5	21	22 AAB51126	Mammaryglobln synthe
39	77	15.5	21	22 AAB51126	Mammaryglobln synthe
40	77	15.5	30	19 AAW48435	Human mammaryglobln
41	72	14.5	20	22 AAB51132	Human mammaryglobln
42	71	14.3	828	22 ABB62649	Drosophila melanog
43	71	14.3	1187	22 ABB63669	Drosophila melanog
44	68.5	13.8	619	22 AAB97025	Human colon carcin
45	68	13.7	55	22 AAB51128	Human mammaryglobln

ALIGNMENTS

RESULT 1	
AAW35804	
ID AAW35804 standard; Protein: 95 AA.	
XX	
AC AAW35804;	
XX	
DT 27-MAR-1998 (first entry)	
XX	
DE Human endometrial specific steroid-binding factor III.	
XX	
KW Endometrial specific steroid-binding factor III; ESF III; human;	
KW Clara cell secretory protein; endometrium;	
KW phospholipase A2 inhibitor; polychlorinated biphenyl; antiagregant;	
KW inflammation; asthma; rhinitis; cystic fibrosis; airway disease;	
KW neoplasia; atopy; therapy; diagnosis.	
XX	
OS Homo sapiens.	
XX	
FH Key	Location/Qualifiers
FT Peptide	1..21
FT Protein	/label= sig_peptide
FT	22..95
FT	/label= Mat_Protein
XX	
PN WO9734997-A1.	
XX	
PD 25-SEP-1997.	
XX	
PF 21-MAR-1996;	96WO-US03857.
XX	
PR 21-MAR-1996;	96WO-US03857.
XX	
PA (HUMA-) HUMAN GENOME SCI INC.	

XX Gentz RL, Ni J, Yu G;
 XX WPI: 1997-480206/44.
 DR N-PSDB: AAT94832.
 XX
 XX Human endometrial specific steroid-binding factor I, II and III -
 PT used to treat inflammation, asthma, rhinitis, cystic fibrosis,
 PT allergy disease, neoplasia, atopy etc.
 XX
 PS Claim 19; Page 65; 92pp; English.
 XX
 CC This sequence comprises human endometrial specific steroid binding
 CC factor III (ESF III), a protein that inhibits phospholipase A2
 CC activity, binds to polychlorinated biphenyl compounds, reduces
 CC foreign protein antigenicity, inhibits monocyte and neutrophil
 CC chemotaxis and phagocytosis, inhibits platelet aggregation,
 CC regulates eicosanoid levels in the human uterus and controls the
 CC growth of endometrial cells. The amino acid sequence was deduced
 CC from a cDNA clone (see AAT94832) derived from a human endometrial
 CC tumor. ESF I (see AAT35802) and ESF II (see AAT35803) are also
 CC claimed. Human ESF III has about 36% identity with rat prostatic
 CC steroid-binding protein. Recombinant ESF I, II and III can be
 CC expressed in host cells for use in claimed methods (a) for treating
 CC a patient in need of ESF I, II or III (including expression of the
 CC polypeptide in vivo) and (b) for identifying compounds which bind
 CC to and inhibit activation of the ESF polypeptide. hESF I, II and
 CC III may be used to treat inflammation, asthma, rhinitis, cystic
 CC fibrosis, allergy disease, neoplasia and atopy.
 XX
 SQ Sequence 95 AA;

Query Match 100.0%; Score 496; DB 18; Length 95;
 Best Local Similarity 100.0%; Pred. No. 4, 5e-48;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLWVLMALALHCHYADSGCKLLEDMVEKTIINSISIPREKELLOEPTIDSDAAAEAMG 60
 DB 1 MKLWVLMALALHCHYADSGCKLLEDMVEKTIINSISIPREKELLOEPTIDSDAAAEAMG 60
 OY 61 KFKQCFLNQSHRTLNKFGIMHNTVYDSIWCNMKSN 95
 DB 61 KFKQCFLNQSHRTLNKFGIMHNTVYDSIWCNMKSN 95

RESULT 2
 ID AAY02590 standard; Protein: 95 AA.

XX AAY02590;
 XX
 DT 26-JUL-1999 (first entry)
 XX
 DE A human mamaglobin homologue (HMH).
 XX
 KW Human mamaglobin homologue; HMH; antagonist; neoplastic disorder;
 KW adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma;
 KW teratocarcinoma; endometritis.
 XX
 OS Homo sapiens.
 XX
 PN W09919487-A1.
 PD 22-APR-1999.
 XX
 PF 14-OCT-1998; 98WO-US21729.
 XX
 PR 16-OCT-1997; 97US-0951750.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Hillman JL, Murry LE, Shah P;
 XX

DR WPI: 1999-302531/25.
 DR N-PSDB: AAX36138.
 XX
 PT New human mamaglobin homolog (HMH), useful for diagnosing, treating
 PT or preventing disorders associated with expression of HMH
 XX
 PS Claim 1; Fig 1A-B; 63pp; English.
 XX

CC The present sequence represents a human mamaglobin homologue (HMH).
 CC Antagonists of the HMH polypeptide can be used to treat neoplastic
 CC disorders including adenocarcinoma, leukemia, lymphoma, melanoma,
 CC myeloma, sarcoma and teratocarcinoma. A vector expressing the
 CC complement of the polynucleotide encoding HMH may be administered
 CC to a subject to treat or prevent neoplastic disorders or endometritis.
 CC Antibodies which bind HMH may also be used in the diagnosis of
 CC conditions or diseases characterized by expression of HMH, or in assays
 CC to monitor patients being treated with HMH, agonists, antagonists or
 CC inhibitors. Polynucleotides encoding HMH may also be used diagnostically
 CC to detect and quantitate gene expression in biopsied tissues. With
 CC respect to cancer a relatively high amount of transcript may indicate a
 CC predisposition for the development of disease. The nucleic acid sequences
 CC which encode HMH may also be used to generate hybridization probes useful
 CC for mapping the naturally occurring genomic sequence. HMH, and its
 CC fragments/variants can be used for screening libraries of compounds in
 CC drug screening techniques.
 XX

Query Match 100.0%; Score 496; DB 20; Length 95;
 Best Local Similarity 100.0%; Pred. No. 4, 5e-48;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLWVLMALALHCHYADSGCKLLEDMVEKTIINSISIPREKELLOEPTIDSDAAAEAMG 60
 DB 1 MKLWVLMALALHCHYADSGCKLLEDMVEKTIINSISIPREKELLOEPTIDSDAAAEAMG 60
 OY 61 KFKQCFLNQSHRTLNKFGIMHNTVYDSIWCNMKSN 95
 DB 61 KFKQCFLNQSHRTLNKFGIMHNTVYDSIWCNMKSN 95

RESULT 3
 ID AAB03769 standard; Protein: 95 AA.

XX AAB03769;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human endometrial specific steroid-binding factor III protein sequence.
 XX
 KW Endometrial specific steroid-binding factor; human; hESF; inflammation;
 KW asthma; rhinitis; cystic fibrosis; air way disease; neoplasia; atopy;
 KW eicosanoid level regulator; chemotaxis inhibitor; endometrial cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN US066724-A.
 PD 23-MAY-2000.
 XX
 PF 21-MAR-1997; 97US-0821451.
 XX
 PR 21-MAR-1996; 96US-0014724.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Yu G, Gentz R, Ni J;
 XX
 DR WPI: 2000-375600/32.
 DR N-PSDB: AAA59730.
 XX
 PT Novel gene encoding human endometrial specific steroid-binding factor

PR I, II and III which is useful for treating asthma, rhinitis, cystic
 PT fibrosis, airway disease and neoplasia
 PS Claim 1; Fig 3; 36pp; English.
 XX
 CC This invention relates to nucleic acid molecules encoding portions of the
 CC human endometrial specific steroid-binding factors I, II, and III. Also
 CC included in the invention are hESF I, II, and III polypeptide sequences.
 CC The nucleotide sequence exhibit antisense, antiproliferatory,
 CC anti-allergic, and cytostatic properties. The polynucleotides are used in
 CC gene therapy to express hESF I, II and III polypeptides in vivo to treat
 CC and/or prevent inflammation, asthma, rhinitis, cystic fibrosis, air way
 CC disease, neoplasia and atopy. The polynucleotides are also used to
 CC inhibit phospholipase A2 activity, bind polychlorinated biphenyls, reduce
 CC foreign protein antigenicity, inhibit monocyte and neutrophil chemotaxis
 CC and phagocytosis, inhibit platelet aggregation, regulate eicosanoid
 CC levels in the human uterus and control the growth of endometrial cells.
 CC The polynucleotides are also useful for detecting complementary
 CC polynucleotides as a diagnostic reagent. The hESF I, II and III
 CC polynucleotides are used to detect complementary polynucleotides such as
 CC a diagnostic reagent. Detection of a mutated form of hESF I, II and III
 CC associated with a dysfunction will provide a diagnostic tool that can
 CC define diagnosis of a disease or susceptibility to a disease which
 CC results from under-expression, over-expression or altered expression of
 CC hESF I, II and III e.g. a susceptibility to inherited asthma and
 CC endometrial cancer. They are also useful for chromosome identification.
 CC The present sequence represents a hESF III protein sequence identified in
 CC the invention.
 CC
 CC Sequence 95 AA;
 SO
 Query Match 100.0%; Score 496; DB 21; Length 95;
 Best Local Similarity 100.0%; Pred. No. 4.5e-48;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MKLMTVMTLAAALHCHYADSGCKLLEDVWEKTNISDIPEYKELLOEFIDSDAAAEAMG 60
 DB 1 MKLMTVMTLAAALHCHYADSGCKLLEDVWEKTNISDIPEYKELLOEFIDSDAAAEAMG 60
 OY 61 KFKOCFLNOSHRTLKNFGLMHTVYDSIMCNMKS 95
 DB 61 KFKOCFLNOSHRTLKNFGLMHTVYDSIMCNMKS 95
 RESULT 4
 AAY92226
 ID AAY92226 standard; Protein; 95 AA.
 AC AAY92226;
 DT 10-AUG-2000 (first entry)
 DE Human endometrial specific steroid binding factor III.
 XX
 KW Endometrial specific steroid binding factor; ESBPIII; diagnosis;
 KW gynaecological cancer; uterine; breast; endometrial; ovarian;
 KW antibody; gene therapy; cytostatic.
 OS Homo sapiens.
 PN WO200020044-A1.
 PD 13-Apr-2000.
 PF 30-SEP-1999; 99WO-US22753.
 PR 02-OCT-1998; 98US-0102743.
 PA (DIAD-) DIADEXUS LLC.
 PI MacIna RA;
 DR WPI; 2000-303649/26.

DR N-PSDB: AAA09104.
 XX
 CC Diagnosing, staging and monitoring gynecological cancer comprising
 PT using an elevated level of ESBPIII in a patient as an indicator of
 PT cancer
 XX
 PS Claim 6; Page 28-29; 32pp; English.
 XX
 CC The levels of human endometrial specific steroid binding factor
 CC (ESBPIII) can be measured and compared to control levels and used to
 CC diagnose the presence of a gynaecological (uterine, breast, endometrial,
 CC or ovarian) cancer in a patient. ESBPIII levels can also be used to
 CC diagnose metastasis, to stage or monitor gynaecological cancer.
 CC Antibodies specific for ESBPIII can be used to treat gynecological
 CC cancers.
 CC
 CC Sequence 95 AA;
 SO
 Query Match 100.0%; Score 496; DB 21; Length 95;
 Best Local Similarity 100.0%; Pred. No. 4.5e-48;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 MKLMTVMTLAAALHCHYADSGCKLLEDVWEKTNISDIPEYKELLOEFIDSDAAAEAMG 60
 DB 1 MKLMTVMTLAAALHCHYADSGCKLLEDVWEKTNISDIPEYKELLOEFIDSDAAAEAMG 60
 OY 61 KFKOCFLNOSHRTLKNFGLMHTVYDSIMCNMKS 95
 DB 61 KFKOCFLNOSHRTLKNFGLMHTVYDSIMCNMKS 95
 RESULT 5
 AAY92237
 ID AAY92237 standard; Protein; 95 AA.
 AC AAY92237;
 DT 10-AUG-2000 (first entry)
 DE Mammoglobin homologue from clone Mamm-X.
 XX
 KW Clone Mamm-X; mammoglobin; breast cancer; cytostatic; anti-HIV;
 KW immunosuppressive; anti-allergic; anti-infective; anti-inflammatory;
 KW anti-atherosclerotic; anti-arteriosclerotic; vasotrophic; neuroprotective;
 KW nootropic; dermatological; tranquilizer; vulnerary.
 OS Homo sapiens.
 PN WO200020047-A2.
 PD 13-Apr-2000.
 PF 06-OCT-1999; 99WO-US23294.
 PR 06-OCT-1998; 98US-0103195.
 PR 05-OCT-1999; 99US-0103195.
 PA (CURA-) CURAGEN CORP.
 PI Shimkets RA;
 DR WPI; 2000-303741/26.
 DR N-PSDB: AAA09118.
 PT Nucleic acids encoding polypeptides with syncollin-like, claudin-like or
 PT cytokine-like activity, useful for treating diseases including cancer,
 PT Alzheimer's and atherosclerosis
 XX
 PS Claim 23; Fig 6; 118pp; English.
 CC Clone Mamm-X encodes a polypeptide that is 100 percent identical to human
 CC Mammoglobin B precursor, a potential marker of breast cancer nodal
 CC metastasis. The sequences are useful for treatment of diseases such as

CC cancer, immune disorders, autoimmune disease, acquired immune deficiency
 CC syndrome (AIDS), transplant rejection, allergy, infection by a
 CC pathological agent or organism, inflammatory disorders, arthritis, a
 CC hematopoietic disorder, a skin disorder, atherosclerosis, restenosis, a
 CC neurological disease, Alzheimer's disease, trauma, spinal cord injury
 CC and skeletal disorders.
 CC
 CC
 CC

Sequence 95 AA:

Query Match 100.0% Score 496; DB 21; Length 95;
 Best Local Similarity 100.0%; Pred. No. 4.5e-48;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLWMLAALLHCHYADSGCKLLEDVWEKTIINSISPEYKELQEFIDSDAAAEAMG 60
 Db 1 MKLWMLAALLHCHYADSGCKLLEDVWEKTIINSISPEYKELQEFIDSDAAAEAMG 60
 OY 61 KFKOCFLNOSHRTLKNFGMLMHTVYDSIMCNKSN 95
 Db 61 KFKOCFLNOSHRTLKNFGMLMHTVYDSIMCNKSN 95

RESULT 6
 ID AAY65394 standard; Protein; 95 AA.
 AAY65394:

01-FEB-2000 (first entry)

Human 5' EST related polypeptide SEQ ID NO:1555.

Human: 5' EST; expressed sequence tag; secreted protein; diagnosis;
 forensic; chromosome mapping; upstream regulatory sequence;
 regulation; identification.

Homo sapiens.

MO9953051-A2.

21-OCT-1999.

09-APR-1999: 99NC-IB00712.

09-APR-1998: 98US-0057719.

28-APR-1998: 98US-0069047.

(GEST) GENSET.

Dumas Milne Edwards J, Duclert A, Giordano J;

WPI: 2000-038446/03.

N-PSDB: AA243008.

Novel secreted protein 5' expressed sequence tag sequences used in
 diagnostic, forensic, gene therapy, and chromosome mapping procedures

Claim 3; Page 818: 837pp; English.

AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)
 sequences, corresponding to human secreted proteins. AA164651 to
 AA165338 represent the EST-related proteins corresponding to AA242265 to
 AA243052. The 5' ESTs can be used for producing secreted human gene
 products. They can be used to identify and isolate 5' untranslated
 regions (UTRs) and upstream regulatory regions which control the
 location, development stage, rate, and quantity of protein synthesis, as
 well as stability of mRNA. The ESTs are also useful as probes for
 chromosome mapping, and to obtain full length cDNA clones. The ESTs can
 also be used in forensic procedures to identify individuals, or in
 diagnostic procedures to identify individuals having genetic diseases
 resulting from abnormal gene expression. The products may also be used in
 gene therapy protocols. The nucleic acids encoding signal peptides can be

CC used for directing extracellular secretion of a polypeptide or the
 CC insertion of a polypeptide into a membrane, or importing a polypeptide
 CC into a cell. The proteins encoded by the EST sequences may be useful in
 CC treating a variety of human conditions. Secreted proteins have
 CC therapeutic value, and the identification of new secreted proteins is
 CC valuable. AA242249 to AA242264 and AA164644 to AA164650 represent
 CC sequences used in the exemplification of the present invention.
 CC
 CC
 CC

Sequence 95 AA:

Query Match 100.0% Score 496; DB 21; Length 95;
 Best Local Similarity 100.0%; Pred. No. 4.5e-48;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLWMLAALLHCHYADSGCKLLEDVWEKTIINSISPEYKELQEFIDSDAAAEAMG 60
 Db 1 MKLWMLAALLHCHYADSGCKLLEDVWEKTIINSISPEYKELQEFIDSDAAAEAMG 60
 OY 61 KFKOCFLNOSHRTLKNFGMLMHTVYDSIMCNKSN 95
 Db 61 KFKOCFLNOSHRTLKNFGMLMHTVYDSIMCNKSN 95

RESULT 7
 ID AAB31682 standard; Protein; 95 AA.
 AAB31682:

30-APR-2001 (first entry)

An endometrial specific steroid binding factor III.

Human: endometrial specific steroid binding factor; hESF; hESF1; hESF11;
 hESF11; inflammation; asthma; rhinitis; cystic fibrosis; airway disease;
 neoplasia; atopy; phospholipase A2; polychlorinated biphenyl; chemotaxis;
 phagocytosis; platelet aggregation; eicosanoid; endometrial cell.

Homo sapiens.

US6174992-B1.

16-JAN-2001.

08-MAR-1999: 99US-0263810.

21-MAR-1996: 96US-0014724.

21-MAR-1997: 97US-0821451.

(HUMA-) HUMAN GENOME SCI INC.

Ni J, Yu G, Gentz R;

WPI: 2001-158477/16.

N-PSDB: AAF25214.

New human endometrial specific steroid binding factors, useful for
 treating and preventing inflammation, asthma, rhinitis, cystic
 fibrosis, airway disease, neoplasia and atopy

Claim 1; Fig 3: 36pp; English.

The present sequence represents a human endometrial specific steroid
 binding factor (hESF). The specification describes hESF1, hESF11, and
 hESF11. hESF1, II and III polypeptides, and polynucleotides encoding
 them are useful for treating and preventing inflammation, asthma,
 rhinitis, cystic fibrosis, airway disease, neoplasia and atopy,
 inhibiting phospholipase A2 activity, binding polychlorinated
 biphenyls, reducing foreign protein antigenicity, inhibiting monocyte

CC and neutrophil chemotaxis and phagocytosis, inhibiting platelet
CC aggregation, regulating eicosanoid levels in the human uterus, and for
CC controlling the growth of endometrial cells. hESF polypeptides and
CC nucleotides are also useful for research, biological, clinical or
CC therapeutic purposes.

XX
SQ Sequence 95 AA:

Query Match 100.0%; Score 496; DB 23; Length 95;
Best Local Similarity 100.0%; Pred. No. 4.5e-48;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLWVLMALALLHCYADSGCKLEDVMEKTIINSISIPYKELLOEPIFSDAAAEAMG 60
DB 1 MKLWVLMALALLHCYADSGCKLEDVMEKTIINSISIPYKELLOEPIFSDAAAEAMG 60
OY 61 KFKOCFLNOSHRTLNFGIMHMYDSTWCMKSN 95
DB 61 KFKOCFLNOSHRTLNFGIMHMYDSTWCMKSN 95

RESULT 8
AA020555

XX ID AA020555 standard; Protein; 95 AA.

XX AC AA020555;

XX DT 27-JUN-2002 (first entry)

XX DE Protein of human Lipophilin C.

XX KW Immunogenic epitope; hormonally regulated organ; malignant tumour;

XX LM Lipophilin; human.

XX OS Homo sapiens.

XX PN US2002034739-A1.

XX PD 21-MAR-2002.

XX PF 07-JUL-1998; 98US-0110716.

XX PR 07-JUL-1998; 98US-0110716.

XX PA (LEHR/) LEHRER R I.

XX PA (ZHAO/) ZHAO C.

XX PA (GLAS/) GLASGOW B J.

XX PI Lehrer RI, Zhao C, Glasgow BJ;

XX DR WPI: 2002-338922/37.

XX DR N-PSDB: AAK99492.

XX PT Peptides having the sequence of human lipophilin A, B and C are
XX associated with carcinomas of hormonally regulated organs and are
XX useful in the diagnosis and prognosis of various cancers -

XX PS Claim 7; Fig 6; 22pp; English.

CC The invention relates to a peptide comprising the amino acid sequences of
CC human lipophilin A, B, or C or its allelic variant or fragment comprising
CC at least one immunogenic epitope, which is purified and isolated, and may
CC have the N-terminal acetylated and/or C-terminal amidated or be a fusion
CC protein. Molecules of the invention are used in the diagnosis and
CC prognosis of malignant tumours associated with hormonally regulated
CC organs such as uterus, ovary, prostate, testis, breast, kidney and
CC thymus. This sequence represents the human lipophilin C protein of the
CC invention.

XX SQ Sequence 95 AA:

Query Match 100.0%; Score 496; DB 23; Length 95;
Best Local Similarity 100.0%; Pred. No. 4.5e-48;

Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLWVLMALALLHCYADSGCKLEDVMEKTIINSISIPYKELLOEPIFSDAAAEAMG 60
DB 1 MKLWVLMALALLHCYADSGCKLEDVMEKTIINSISIPYKELLOEPIFSDAAAEAMG 60

OY 61 KFKOCFLNOSHRTLNFGIMHMYDSTWCMKSN 95
DB 61 KFKOCFLNOSHRTLNFGIMHMYDSTWCMKSN 95

RESULT 9
AB09635

XX ID AB09635 standard; Protein; 95 AA.

XX AC AB09635;

XX DT 29-MAR-2002 (first entry)

XX DE Human endometrial specific steroid-binding factor (hesf) III.

XX KW Human; endometrial specific steroid-binding factor; ESF;

XX KW prostatic steroid-binding protein; hesf I; hesf II; hesf III; asthma.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Peptide 1..21

XX FT /note= "signal peptide"

XX FT Protein 22..95

XX FT /note= "mature protein"

XX PN US6338948-B1.

XX PD 15-JAN-2002.

XX PF 30-MAY-2000; 2000US-0583169.

XX PR 21-MAR-1996; 96US-014724P.

XX PR 21-MAR-1997; 97US-0821451.

XX PR 08-MAR-1999; 99US-0263810.

XX PA (HDNA-) HUMAN GENOME SCI INC.

XX PI NI J, YU G, Gentz R;

XX DR WPI: 2002-215019/27.

XX DR N-PSDB: ABL41783.

XX PT New antibody specific for human endometrial specific steroid-binding
XX factor (hesf) III, useful for detecting hesf III protein in biological
XX sample and to isolate or identify clones expressing the protein -

XX PS Disclosure; Fig 3; 36pp; English.

CC The present sequence represents a endometrial specific steroid-binding
CC factor (hesf) III. The full length protein has a molecular weight of
CC 8.10 kDa. The protein has homology to rat prostatic steroid-binding
CC protein C3. Antibodies which bind hesf proteins, such as hesf I, hesf II,
CC and hesf III are useful for isolating or to identify clones expressing
CC the polypeptides or to purify the polypeptides by affinity
CC chromatography. Agonists and antagonists of hesf proteins are useful
CC for treating and/or preventing susceptibility to asthma.

XX SQ Sequence 95 AA:

Query Match 100.0%; Score 496; DB 23; Length 95;
Best Local Similarity 100.0%; Pred. No. 4.5e-48;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLWVLMALALLHCYADSGCKLEDVMEKTIINSISIPYKELLOEPIFSDAAAEAMG 60
DB 1 MKLWVLMALALLHCYADSGCKLEDVMEKTIINSISIPYKELLOEPIFSDAAAEAMG 60

OY 61 KFKOCFLNOSHRTLKNGFLMMHTVYDSIMCNKSN 95
 DB 61 KFKOCFLNOSHRTLKNGFLMMHTVYDSIMCNKSN 95

RESULT 10

AAV60038
 ID AAV60038 standard; protein; 108 AA.

XX AAV60038:
 AC AAV60038:

XX 31-JAN-2000 (first entry)
 DT

XX Human endometrium tumour EST encoded protein 98.
 DE

XX Endometrium: human; tumour; cancer; anticancer; cytostatic; EST:
 KW treatment; uterine; gene therapy; expressed sequence tag.

XX Homo sapiens.
 OS

XX DE19817948-A1.
 PN

XX 21-OCT-1999.
 PD

XX 17-APR-1998; 98DE-1017948.
 PP

XX 17-APR-1998; 98DE-1017948.
 PR

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 PA

XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
 PI WPI: 1999-591957/51.

XX N-PSDB; AA242013.
 DR

PT New nucleic acid sequences expressed in uterine cancer tissues, and
 derived polypeptides, for treatment of uterine and endometrial cancer
 and identification of therapeutic agents -

XX Claim 23; Page 314; 444pp; German.

XX This invention describes novel human nucleic acid (cDNA) sequences (A),
 that are highly expressed in uterine tumour tissue and which have
 anticancer and cytostatic activity. (A) are used (i) for recombinant
 expression of polypeptides (B) and (ii) to isolate complete genes. (B)
 are used (i) to identify agents suitable for treatment of uterine or
 endometrial cancer; (ii) directly for treating these forms of cancer
 (including expression from gene therapy vectors) and (iii) for
 generation of specific antibodies. (A) are identified by assembling ESTs
 (expressed sequence tags) from a particular tissue type before comparison
 of expression patterns. This allows a significantly longer fragment of
 the gene to be revealed, so should reduce the number of failures
 associated with the fact that ESTs from different libraries may represent
 different parts of the same unknown gene, distorting the estimated
 frequency of occurrence in a particular tissue. AAV5941-Y60328 represent
 cDNA fragments encoded by the human endometrium tumour cDNA library
 derived EST fragments represented in AA241981-242121.

XX Sequence 108 AA;

Query Match 100.0%; Score 496; DB 20; Length 108;
 Best Local Similarity 100.0%; Pred. No. 5.3e-48;

Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLLMVLALALLHGYADSGCKLLEDMVEKTIINSIPKELLOEFIDSDAAAEAMG 60
 DB 14 MKLLMVLALALLHGYADSGCKLLEDMVEKTIINSIPKELLOEFIDSDAAAEAMG 73

OY 61 KFKOCFLNOSHRTLKNGFLMMHTVYDSIMCNKSN 95
 DB 74 KFKOCFLNOSHRTLKNGFLMMHTVYDSIMCNKSN 108

RESULT 11

AAW61649
 ID AAW61649 standard; peptide; 74 AA.

XX AAW61649;
 AC AAW61649;

XX 27-OCT-1998 (first entry)
 DT

XX Non-ocular disease marker 3.
 DE

XX Human; non-ocular disease; tear; cancer; breast; prostate.
 KW

XX Homo sapiens.
 OS

XX WO9835229-A1.
 PN

XX 13-AUG-1998.
 PD

XX 06-FEB-1998; 98WO-AU00071.
 PP

XX 07-FEB-1997; 97AU-0005009.
 PR

XX (MACO-) MACOUIARIE RES LTD.
 PA (UNIX) UNISEARCH LTD.

XX Bolis S, Gooley AA, Herbert B, Molloy M, Morris C;
 PI Walsh B, Willcox M, Williams KL;

XX WPI: 1998-447373/38.
 DR

XX Screening for non-ocular disease - by analysing tears for marker
 proteins, particularly indicative of cancer and genetic disease.
 PT also new proteins and nucleic acid encoding them

XX Claim 6; Page 9; 14pp; English.

XX The markers AAW61647-W61649 are used for screening for, or detecting,
 CC non-ocular disease by analysing tears. Biochemicals, specifically
 CC proteins, are isolated from tears, particularly by chromatography or
 CC electrophoresis, especially two-dimensional polyacrylamide gel
 CC electrophoresis (2D-PAGE), then detected, e.g. with labelled specific
 CC reagents, in (radio)immunoassay. The method is used to detect cancer,
 CC particularly of breast or prostate, or a genetic disease, in humans or
 CC animals.

XX Sequence 74 AA;

Query Match 78.6%; Score 390; DB 19; Length 74;
 Best Local Similarity 98.6%; Pred. No. 2.6e-36;

Matches 73; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 19 DSGCKLLEDMVEKTIINSIPKELLOEFIDSDAAAEAMGKFKOCFLNOSHRTLKNG 78
 DB 1 DSGCKLLEDMVEKTIINSIPKELLOEFIDSDAAAEAMGKFKOCFLNOSHRTLKNG 60

OY 79 LMMHTVYDSIMCNM 92
 DB 61 LMMHTVYDSIMCNL 74

RESULT 12

AAU83620
 ID AAU83620 standard; protein; 93 AA.

XX AAU83620;
 AC AAU83620;

XX 08-MAY-2002 (first entry)
 DT

XX Human PRO protein, Seq ID No 58.
 DE

XX Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
 KW breast cancer; prostate tumour; rectal tumour; liver tumour;

KM pericyte cell proliferation; chondrocyte cell proliferation;
KM tumour necrosis factor-alpha.
XX Homo sapiens.
XX OS
XX WO200208288-A2.
XX
XX 31-JAN-2002.
XX
XX 29-JUN-2001; 2001WO-US21066.
XX
XX 20-JUL-2000; 2000US-219556P.
XX 25-JUL-2000; 2000US-220585P.
XX 25-JUL-2000; 2000US-220605P.
XX 25-JUL-2000; 2000US-220607P.
XX 25-JUL-2000; 2000US-220624P.
XX 25-JUL-2000; 2000US-220638P.
XX 25-JUL-2000; 2000US-220664P.
XX 25-JUL-2000; 2000US-220666P.
XX 26-JUL-2000; 2000US-220893P.
XX 28-JUL-2000; 2000WO-US20710.
XX 23-AUG-2000; 2000WO-US23352.
XX 24-AUG-2000; 2000WO-US23328.
XX 15-SEP-2000; 2000US-000000P.
XX 10-NOV-2000; 2000WO-US30873.
XX 28-NOV-2000; 2000US-253646P.
XX 01-DEC-2000; 2000WO-US32678.
XX 20-DEC-2000; 2000US-0747259.
XX 20-DEC-2000; 2000WO-US34956.
XX 28-FEB-2001; 2001WO-US06520.
XX 10-MAY-2001; 2001US-0854280.
XX 25-MAY-2001; 2001WO-US17092.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WT;
XX
XX WPI: 2002-172001/22.
XX N-PSDB: ABK33564.
XX
XX One hundred and twenty two nucleic acids encoding PRO polypeptides,
XX useful for treating a PRO related disorder and for diagnosing tumours
XX such as lung cancer, colon cancer, breast tumour, prostate tumour, rectal
XX tumour or liver tumour -
XX
XX Claim 11; Figure 58; 359pp; English.
XX
XX The invention relates to one hundred and twenty two nucleic acids
XX encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides
XX encode human secreted proteins. The PRO nucleic acids, polypeptides,
XX agonists and antagonists are useful for treating a PRO related disorder.
XX The PRO polypeptides are useful for diagnosing tumours, especially lung
XX cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or
XX liver tumour. The PRO polypeptides are useful for stimulating the
XX proliferation of, or gene expression, in pericyte cells, for stimulating
XX the proliferation or differentiation of chondrocyte cells, for
XX stimulating the release of tumour necrosis factor-alpha from human blood,
XX for stimulating or inhibiting the proliferation of normal human dermal
XX fibroblast cells. The PRO polypeptide may also be used as molecular
XX weight markers and for tissue typing. The PRO nucleic acids have
XX applications in molecular biology, including use as hybridisation probes,
XX and in chromosome and gene mapping. AAU83592-AAU83713 represent human PRO
XX protein sequences of the invention.
XX
XX Sequence 93 AA:
XX
XX Query Match 57.18; Score 283; DB 23; Length 93;
XX Best Local Similarity 57.88; Pred. No. 3.6e-24;
XX Matches 52; Conservative 20; Mismatches 18; Indels 0; Gaps 0;
XX
XX 1 MKLLVLMIALLLHCHYDSCGCKLEDMVEKTNISDIPEYKELLQEFIDSDAAAEAMG 60
XX :::::|||||:||||| ||| ||| ::|||:: ::||| || |||::: |

DB 1 MKVWVLLMLAALPLCYAGSGCVLLESVEKTIIDPSVVEFKADLRIDTEQEAWE 60
QY 61 KFKCCFLNQSHRTLKNGFLMHTVYDSIMC 90
:::||||:||||| ||| ||| ::|||:||||:|
DB 61 EFKECFLSQSNETLANFRVWHTIYDSLXC 90
RESULT 13
AAW10179
ID AAW10179 standard; Protein: 93 AA.
XX
XX AAW10179;
XX
XX 12-AUG-1997 (first entry)
XX
XX Mammary-specific secretory protein, mammaglobin.
XX
XX mammaglobin; mammary-secretory protein; breast cancer; detection;
XX neoplastic disease; diagnosis.
XX
XX Homo sapiens.
XX
XX OS
XX FH Key location/Qualifiers
XX FT Peptide 1..19
XX FT /label= signal-peptide
XX FT Protein 20..93
XX /label= mature_protein
XX
XX W09638463-A1.
XX
XX 05-DEC-1996.
XX
XX PD 31-MAY-1996; 96WO-US08235.
XX
XX 31-MAY-1995; 95US-0455896.
XX
XX PA (UNITM) UNIV WASHINGTON.
XX
XX PI Fleming TP, Watson MA;
XX
XX WPI: 1997-034259/03.
XX N-PSDB: AAT50925.
XX
XX Nucleic acid encoding mammary-specific secretory protein,
XX mammaglobin - used to develop prods. for the early diagnosis and
XX treatment of breast cancer neoplastic disease
XX
XX Claim 3; Fig 2; 54pp; English.
XX
XX The present sequence is that of a mammary-specific secretory protein
XX designated mammaglobin, which is overexpressed in 27% of stage I primary
XX breast cancer tumours. The anonymous sequence lay previously designated
XX DE87002 was used to demonstrate that mammaglobin is abundant in the
XX breast cancer tumour cell line MDA-MB-415. To isolate the full-length
XX mammaglobin cDNA (AAT50925), the mRNA was reverse transcribed from
XX this cell line and cloned using the RACE PCR technique. The nucleic acid
XX and protein can be used to develop prods. e.g. antibodies or probes, for
XX the detection and treatment of breast neoplastic disease.
XX
XX Sequence 93 AA:
XX
XX Query Match 56.7%; Score 281; DB 18; Length 93;
XX Best Local Similarity 58.7%; Pred. No. 6e-24;
XX Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;
XX
XX 1 MKLLVLMIALLLHCHYDSCGCKLEDMVEKTNISDIPEYKELLQEFIDSDAAAEAMG 60
XX ::|||:||||| ||| ||| ::|||:||||:|
DB 61 KFKCCFLNQSHRTLKNGFLMHTVYDSIMC 92
XX :|||||: ||| ||| ::|||:||||:|
DB 61 ELKECFLNGTDETLNSVVEFMQLIYDSSLCDL 92

```

RESULT 14
ID AAM59777 standard; Protein: 93 AA.
XX
AC AAM59777:
XX
DT 12-OCT-1998 (first entry)
XX
DE Amino acid sequence of the human steroid binding protein C2.
XX
KW Human steroid-binding protein C2; hSBP2; hSBP1; breast cancer; probe;
KW gene therapy vector; ribozyme; probe; hybridisation; amplification;
KW antibody; immunoassay.
XX
OS Homo sapiens.
XX
PN WP0821331-A1.
XX
PD 22-MAY-1998.
XX
PE 07-NOV-1997; 97MO-US20674.
XX
PR 12-NOV-1996; 96US-0747547.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Akerblom IE, Goll SK, Hawkins PR, Hillman JL, Murry LE:
XX
DR WPI: 1998-297935/26.
XX
DR N-PSDB; AAM41580.
XX
PT New human steroid binding proteins C1 and C2 - useful for, e.g.
PT diagnosis, monitoring and treating breast cancer, and for drug
PT screening
XX
XX
XX Claim 12; Fig 2; 70pp; English.
XX
XX
XX This is the amino acid sequence of the human steroid-binding protein
XX C2 (hSBP2) used in the method of the invention for the diagnosis,
XX monitoring and treatment of breast cancer. hSBP1 and hSBP2 are useful
XX as markers for breast cancer, i.e. measuring levels of hSBP1 and hSBP2
XX used for diagnosis or monitoring the disease, to identify subjects
XX at risk and to discriminate between different forms of cancer for
XX selection of appropriate therapies. They may also be used for drug
XX screening. Nucleic acids encoding hSBP1 and hSBP2 can be used in gene
XX therapy vectors to over express the steroid-binding proteins, preventing
XX binding of steroids, or antisense sequences, ribozymes. Their nucleic
XX acids can also be used for the diagnosis and monitoring (by quantifying
XX expression of hSBP), as source of probes for hybridisation and
XX amplification and for mapping the genomic sequence. Antibodies are used
XX as diagnostic reagents in standard immunoassays for hSBP.
XX
XX
XX Sequence 93 AA:
XX
XX
XX Query Match 56.7%; Score 281; DB 19; Length 93;
XX Best Local Similarity 58.7%; Pred. No. 6e-24;
XX Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;
XX
XX 1 MKLMTVMTLALMLNCTADSGCKLEDNVEKTIINSIDIPYKELDFPDSDAAEAMG 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 1 MKLMTVMTLALMLNCTADSGCKLEDNVEKTIINSIDIPYKELDFPDSDAAEAMG 60
XX
XX 61 KFKOCFLNOSHRTLKNFCGLMHTVYDSIWCNM 92
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 ELKRCFLNCTDETLNVEFMQIYDSSLCDL 92
XX
XX
XX RESULT 15
XX AAM48432
XX ID AAM48432 standard; Protein: 93 AA.
XX

```

[illegible]

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2003, 11:38:34 ; Search time 45 Seconds

(without alignments)
62.115 Million cell updates/sec

Title: US-09-806-302a-2

Sequence: 1 MKLWVLMALILHCYADS.....NEGLMHTVDSIWCNKN 95

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	496	100.0	95	3	US-08-821-451A-6
2	496	100.0	95	4	US-09-263-810-6
3	496	100.0	95	4	US-09-583-169-6
4	281	56.7	93	1	US-08-455-896-2
5	281	56.7	93	2	US-08-933-148-2
6	281	56.7	93	2	US-09-082-343-2
7	281	56.7	93	3	US-09-082-253-2
8	281	56.7	93	4	US-09-215-818-5
9	281	56.7	93	5	PCIT-US96-08235-2
10	196.5	39.6	95	3	US-08-821-451A-27
11	196.5	39.6	95	4	US-09-263-810-27
12	196.5	39.6	95	4	US-09-583-169-27
13	195.5	39.4	95	1	US-08-455-896-7
14	195.5	39.4	95	2	US-08-933-148-7
15	195.5	39.4	95	2	US-09-082-343-7
16	195.5	39.4	95	3	US-09-082-253-7
17	195.5	39.4	95	5	PCIT-US96-08235-7
18	65	13.1	1255	3	US-08-947-823-3
19	63.5	12.8	582	1	US-08-431-080-16
20	63.5	12.8	582	2	US-08-938-534-16
21	63.5	12.8	582	4	US-09-345-294-16
22	61.5	12.4	1604	4	US-09-004-838-95
23	61	12.3	446	2	US-08-822-171-3
24	61	12.3	472	2	US-08-822-171-2
25	59.5	12.0	1257	3	US-08-947-823-5
26	59.5	12.0	114	1	US-08-031-399-3
27	59.5	12.0	114	1	US-08-031-399-6

28	59.5	12.0	114	1	US-08-031-399-12	Sequence 12, Appl
29	59.5	12.0	114	1	US-08-393-305-3	Sequence 3, Appl
30	59.5	12.0	114	1	US-08-393-305-6	Sequence 6, Appl
31	59.5	12.0	114	1	US-08-726-817-3	Sequence 3, Appl
32	59.5	12.0	114	1	US-08-726-817-6	Sequence 6, Appl
33	59.5	12.0	114	1	US-08-504-042-3	Sequence 3, Appl
34	59.5	12.0	114	1	US-08-504-042-6	Sequence 6, Appl
35	59.5	12.0	114	1	US-08-504-042-12	Sequence 12, Appl
36	59.5	12.0	114	2	US-08-725-969-3	Sequence 3, Appl
37	59.5	12.0	114	2	US-08-725-969-6	Sequence 6, Appl
38	59.5	12.0	114	2	US-08-794-524-3	Sequence 3, Appl
39	59.5	12.0	114	2	US-08-794-524-6	Sequence 6, Appl
40	59.5	12.0	114	4	US-09-189-193-3	Sequence 3, Appl
41	59.5	12.0	114	4	US-09-189-193-6	Sequence 6, Appl
42	59.5	12.0	114	5	PCIT-US94-03793-3	Sequence 3, Appl
43	59.5	12.0	114	5	PCIT-US94-03793-6	Sequence 6, Appl
44	59.5	12.0	114	5	PCIT-US94-03793-12	Sequence 12, Appl
45	59.5	12.0	122	1	US-08-300-903A-3	Sequence 3, Appl

ALIGNMENTS

```
RESULT 1
US-08-821-451A-6
? Sequence 6, Application US/08821451A
? Patent No. 6066724
? GENERAL INFORMATION:
? APPLICANT: Jian Ni, Guo-Liang Yu and Reiner Gentz
? TITLE OF INVENTION: Human Endometrial Specific Steroid-
? TITLE OF INVENTION: Binding Factor I, II and III
? NUMBER OF SEQUENCES: 27
? CORRESPONDENCE ADDRESS:
? ADDRESSER: CARELIA, BYRNE, BAIN, GILFILLAN,
? ADDRESSEE: CECCHI, STEWART & OLSTEIN
? STREET: 6 BECKER FARM ROAD
? CITY: ROSELAND
? STATE: NEW JERSEY
? COUNTRY: USA
? ZIP: 07068
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 INCH DISKETTE
? COMPUTER: IBM PS/2
? OPERATING SYSTEM: MS-DOS
? SOFTWARE: WORD PERFECT 5.1
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/821,451A
? FILING DATE: March 21, 1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 60/014,724
? FILING DATE: March 21, 1996
? ATTORNEY/AGENT INFORMATION:
? NAME: MULHINS, J.G.
? REGISTRATION NUMBER: 33,073
? REFERENCE/DOCKET NUMBER: 325800-521 (P257)
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 201-994-1700
? TELEFAX: 201-994-1744
? INFORMATION FOR SEQ ID NO: 6:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 95 AMINO ACIDS
? TYPE: AMINO ACID
? STRANDEDNESS:
? TOPOLOGY: LINEAR
? MOLECULE TYPE: PROTEIN
? US-08-821-451A-6
?
? Query Match 100.0%; Score 496; DB 3; Length 95;
? Best Local Similarity 100.0%; Pred. No. 7,2e+54;
? Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
?
? 1 MKLWVLMALILHCYADSKLLMDWVEKTIINSIDIPKKELQETIDSDAAAEAMG 60
```

```
Db 1 MKLWVLMALALLHCHYADSGCKLLEDMWEKTTNSDISIPEYKELLOEFISDAAAEAMG 60
QY 61 KFKOCFLNOSHRTLKNFGLMMHTVYDSIMCNMKS 95
Db 61 KFKOCFLNOSHRTLKNFGLMMHTVYDSIMCNMKS 95

RESULT 2
US-09-263-810-6
; Sequence 6, Application US/09263810
; Patent No. 6174992
; GENERAL INFORMATION:
; APPLICANT: JIAN NI, GUO-LIANG YU AND REINER GENTZ
; TITLE OF INVENTION: Human Endometrial Specific Steroid-
; TITLE OF INVENTION: Binding Factor I, II and III
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/263,810
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/821,451
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325600-521 (PF257)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-09-263-810-6

Query Match 100.0%; Score 496; DB 4; Length 95;
Best Local Similarity 100.0%; Pred. No. 7.2e-54;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
; TITLE OF INVENTION: Binding Factor I, II and III
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/583,169
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/821,451
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325600-521 (PF257)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-09-583-169-6

Query Match 100.0%; Score 496; DB 4; Length 95;
Best Local Similarity 100.0%; Pred. No. 7.2e-54;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MKLWVLMALALLHCHYADSGCKLLEDMWEKTTNSDISIPEYKELLOEFISDAAAEAMG 60
QY 61 KFKOCFLNOSHRTLKNFGLMMHTVYDSIMCNMKS 95
Db 61 KFKOCFLNOSHRTLKNFGLMMHTVYDSIMCNMKS 95

RESULT 4
US-08-455-896-2
; Sequence 2, Application US/08455896
; Patent No. 5668267
; GENERAL INFORMATION:
; APPLICANT: HATSON, MARK A.
; APPLICANT: FLEMING, TIMOTHY P.
; TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
; TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: ROGERS, HOWELL & HAFERKAMP
; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
; CITY: ST. LOUIS
; STATE: MISSOURI
; COUNTRY: USA
; ZIP: 63105-1817
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/455,896
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: HOLLAND, DONALD R.
: REGISTRATION NUMBER: 35,197
: REFERENCE/DOCKET NUMBER: 952726
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (314) 727-6092
: TELEFAX: (314) 727-6092
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 93 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: NO
: US-08-455-896-2

Query Match
Best Local Similarity 56.7%; Score 281; DB 1; Length 93;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

OY 1 MRLVLMALALLHCYADSGCKLEDMVEKTIINSISPEYKELLOEFIDSDAAEAMG 60
DB 1 MRLVLMALALSOHCYAGSGCPLELVISKTIINPOVSKTEYKELLOEFIDNATTNAID 60
OY 61 KFKOCFLNOSHRTLNFGMLMHTVYDSINCMN 92
DB 61 ELKECFLNQTDETLSNVEVFMQLYDSSLCDL 92

RESULT 5
US-08-933-149-2
: Sequence 2, Application US/08933149
: Patent No. 5922836
: GENERAL INFORMATION:
: APPLICANT: WATSON, MARK A.
: APPLICANT: FLEMING, TIMOTHY P.
: TITLE OF INVENTION: MAMMAGLOBIN, A SECRETED
: TITLE OF INVENTION: MAMMARY SPECIFIC BREAST CANCER PROTEIN
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESS: HOWELL & HAERKAMP, L.C.
: STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
: CITY: ST. LOUIS
: STATE: MISSOURI
: COUNTRY: USA
: ZIP: 63105-1817
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/933,149
: FILING DATE:
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: HENDERSON, MELODIE W.
: REGISTRATION NUMBER: 37,848
: REFERENCE/DOCKET NUMBER: 6029-6040
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (314) 727-5188
: TELEFAX: (314) 727-6092
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 93 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
```

```

: MOLECULE TYPE: protein
: HYPOTHEICAL: NO
: US-08-933-149-2

Query Match
Best Local Similarity 56.7%; Score 281; DB 2; Length 93;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

OY 1 MRLVLMALALLHCYADSGCKLEDMVEKTIINSISPEYKELLOEFIDSDAAEAMG 60
DB 1 MRLVLMALALSOHCYAGSGCPLELVISKTIINPOVSKTEYKELLOEFIDNATTNAID 60
OY 61 KFKOCFLNOSHRTLNFGMLMHTVYDSINCMN 92
DB 61 ELKECFLNQTDETLSNVEVFMQLYDSSLCDL 92

RESULT 6
US-09-082-343-2
: Sequence 2, Application US/09082343
: Patent No. 5968754
: GENERAL INFORMATION:
: APPLICANT: WATSON, MARK A.
: APPLICANT: FLEMING, TIMOTHY P.
: TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
: TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESS: ROGERS, HOWELL & HAERKAMP
: STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
: CITY: ST. LOUIS
: STATE: MISSOURI
: COUNTRY: USA
: ZIP: 63105-1817
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/082,343
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/455,896
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: HOLLAND, DONALD R.
: REGISTRATION NUMBER: 35,197
: REFERENCE/DOCKET NUMBER: 952726
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (314) 727-5188
: TELEFAX: (314) 727-6092
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 93 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: NO
: US-09-082-343-2

Query Match
Best Local Similarity 56.7%; Score 281; DB 2; Length 93;
Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

OY 1 MRLVLMALALLHCYADSGCKLEDMVEKTIINSISPEYKELLOEFIDSDAAEAMG 60
DB 1 MRLVLMALALSOHCYAGSGCPLELVISKTIINPOVSKTEYKELLOEFIDNATTNAID 60
OY 61 KFKOCFLNOSHRTLNFGMLMHTVYDSINCMN 92
DB 61 ELKECFLNQTDETLSNVEVFMQLYDSSLCDL 92
```

Db 61 ELKECFLNQDETLSNVEFMQLIYDSSLCDL 92

RESULT 7

US-09-082-253-2

; Sequence 2, Application US/09082253

; Patent No. 6004756

; GENERAL INFORMATION:

; APPLICANT: MATSON, MARK A.

; APPLICANT: FLEMING, TIMOTHY P.

; TITLE OF INVENTION: DNA SEQUENCE AND ENCODED

; MAMMARY-SPECIFIC BREAST CANCER PROTEIN

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ROGERS, HOWELL & HAFERKAMP

; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400

; CITY: ST. LOUIS

; STATE: MISSOURI

; COUNTRY: USA

; ZIP: 63105-1817

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/082,253

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/455,896

; FILING DATE: 05/31/1995

; ATTORNEY/AGENT INFORMATION:

; NAME: HOLLAND, DONALD R.

; REGISTRATION NUMBER: 35,197

; REFERENCE/DOCKET NUMBER: 952726

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (314) 727-5188

; TELEFAX: (314) 727-6092

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 93 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHEICAL: NO

US-09-082-253-2

Query Match 56.7%; Score 281; DB 3; Length 93;

Best Local Similarity 58.7%; Pred. No. 2, 2e-27;

Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

Qy 1 MRLWVLMALALHCHYADSGCKLLEDMVEKTIINSISIPYEKELLOEFIDSDAAAEAMG 60

Db 1 MRLWVLMALALHCHYADSGCKLLEDMVEKTIINSISIPYEKELLOEFIDSDAAAEAMG 60

Qy 61 KEKCEFLNQSHTLKNFGMLMHTVYDSINCM 92

Db 61 KEKCEFLNQSHTLKNFGMLMHTVYDSINCM 92

Qy 61 ELKECFLNQDETLSNVEFMQLIYDSSLCDL 92

Db 61 ELKECFLNQDETLSNVEFMQLIYDSSLCDL 92

RESULT 8

US-09-215-818-5

; Sequence 5, Application US/09215818A

; Patent No. 6379671

; GENERAL INFORMATION:

; APPLICANT: COLPITS, Tracey

; TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR

; DETECTING DISEASES OF THE BREAST

; FILE REFERENCE: 5972, US, P2

; CURRENT APPLICATION NUMBER: US/09/215,818A.

; CURRENT FILING DATE: 1998-12-18

; EARLIER APPLICATION NUMBER: 08/912,276

; EARLIER FILING DATE: 1997-08-17

; EARLIER APPLICATION NUMBER: 08/697,105

; EARLIER FILING DATE: 1996-08-19

; EARLIER APPLICATION NUMBER: 08/912,149

; EARLIER FILING DATE: 1997-08-15

; EARLIER APPLICATION NUMBER: 08/697,106

; EARLIER FILING DATE: 1996-08-19

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: Pasted for Windows Version 3.0

; SEQ ID NO 5

; LENGTH: 93

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-215-818-5

Query Match 56.7%; Score 281; DB 4; Length 93;

Best Local Similarity 58.7%; Pred. No. 2, 2e-27;

Matches 54; Conservative 13; Mismatches 25; Indels 0; Gaps 0;

Qy 1 MRLWVLMALALHCHYADSGCKLLEDMVEKTIINSISIPYEKELLOEFIDSDAAAEAMG 60

Db 1 MRLWVLMALALHCHYADSGCKLLEDMVEKTIINSISIPYEKELLOEFIDSDAAAEAMG 60

Qy 61 KEKCEFLNQSHTLKNFGMLMHTVYDSINCM 92

Db 61 KEKCEFLNQSHTLKNFGMLMHTVYDSINCM 92

Qy 61 ELKECFLNQDETLSNVEFMQLIYDSSLCDL 92

Db 61 ELKECFLNQDETLSNVEFMQLIYDSSLCDL 92

RESULT 9

PCT-US96-08235-2

; Sequence 2, Application PC/TUS9608235

; GENERAL INFORMATION:

; APPLICANT: MATSON, MARK A.

; APPLICANT: FLEMING, TIMOTHY P.

; TITLE OF INVENTION: DNA SEQUENCE AND ENCODED

; MAMMARY-SPECIFIC BREAST CANCER PROTEIN

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ROGERS, HOWELL & HAFERKAMP

; STREET: 7733 FORSYTH BOULEVARD, SUITE 1400

; CITY: ST. LOUIS

; STATE: MISSOURI

; COUNTRY: USA

; ZIP: 63105-1817

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/08235

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: HOLLAND, DONALD R.

; REGISTRATION NUMBER: 35,197

; REFERENCE/DOCKET NUMBER: 964796

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (314) 727-5188

; TELEFAX: (314) 727-6092

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 93 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHEICAL: NO

PCT-US96-08235-2

Query Match 56.7%; Score 281; DB 5; Length 93;

Best Local Similarity 58.7%; Pred. No. 2, 2e-27;

GENERAL INFORMATION:
APPLICANT: MATSON, MARK A.
APPLICANT: FLEMING, TIMOTHY P.
TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
TITLE OF INVENTION: MAAMARY-SPECIFIC BREAST CANCER PROTEIN
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROGERS, HOWELL & HAFERKAMP
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: USA
ZIP: 63105-1817
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,896
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 95/726
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 7:

```
Qy      1 MKLMLVLMIAILLHCYAD-SGCKLEDMVEKTIINSIDISIPXKELLQEIFSDAAAEAM 59
      |||:::|:::|||||:::||||:::|::::::
Db      1 MKLVLEFLVLTIPICCAYSGSGCSLDEVTGRTINSTVTLHDMKLVKPRVQDHFTEKAV 60
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2003, 13:26:19 : Search time 57 seconds
(without alignments) 160.224 Million cell updates/sec

Title: US-09-806-302a-2

Perfect score: 496

Sequence: 1 MKLMLVLMALILHICVADS.....NFGMLMHTVYDSIMCNKSN 95

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: PIR-73:*
2: p1r1:*
3: p1r2:*
4: p1r3:*
5: p1r4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	195.5	39.4	95	1 BORT3	prostatic steroid-
2	76.5	15.4	280	2 T38449	hypothetical prote
3	76	15.3	212	2 H87530	D-alanyl-D-alanine
4	73.5	14.8	590	2 A96556	unknown protein, 3
5	71.5	14.4	131	2 F70880	hypothetical prote
6	71	14.3	303	2 E97850	cell division prot
7	70	14.1	98	1 BORT2	prostatic steroid-
8	70	14.1	303	2 B71538	cell division prot
9	69.5	14.0	412	2 T47321	hypothetical prote
10	69.5	14.0	1048	2 C96659	protein FIN19.15 [
11	68.5	13.8	173	2 T01889	hypothetical prote
12	68.5	13.8	213	2 C96925	probable membrane
13	68.5	13.8	219	2 F69435	hypothetical prote
14	68.5	13.8	496	2 T46356	hypothetical prote
15	68	13.7	516	2 H87369	hypothetical prote
16	67.5	13.6	268	1 GWKEC	tryptophan halogen
17	67	13.5	457	2 T29741	indole-3-glycerol-
18	67	13.5	738	2 C84700	hypothetical prote
19	66.5	13.4	95	2 S68231	hypothetical prote
20	66.5	13.4	234	2 F64104	PH622 protein prec
21	66	13.3	638	2 T47569	hypothetical prote
22	65.5	13.2	162	2 T49124	hypothetical prote
23	65.5	13.2	429	2 D84554	interleukin-15 - m
24	65.5	13.2	603	1 W1M18	probable cyclin 2
25	65	13.1	269	1 T04394	EB1 protein - human
26	65	13.1	975	2 A86258	NBS-LRR type resist
27	65	13.1	1255	2 T06267	protein F5011.4 [1
28	65	13.1	2145	2 S61041	nematodes resistan
29	64.5	13.0	153	2 E82928	glutamate synthase
					hypothetical prote

30	64.5	13.0	288	2 H84757	hypothetical prote
31	64.5	13.0	490	2 T31646	hypothetical prote
32	64.5	13.0	610	2 S59558	dynamun-like prote
33	64	12.9	629	2 A96657	unknown protein, 7
34	63.5	12.8	264	2 G69278	glutamine ABC tran
35	63.5	12.8	269	2 G75148	hypothetical prote
36	63.5	12.8	582	2 S69720	hypothetical prote
37	63.5	12.8	1294	2 S58307	hypothetical prote
38	63	12.7	131	2 D81908	hypothetical prote
39	63	12.7	288	2 AB0092	hypothetical prote
40	63	12.7	785	2 D7153	probable flagellar
41	62.5	12.6	91	1 UGRB	cell surface antiq
42	62.5	12.6	113	2 JC4143	uteroglobin precu
43	62.5	12.6	137	2 F97763	molt-inhibiting ho
44	62.5	12.6	223	2 S50404	hypothetical prote
45	62.5	12.6	429	2 T04667	hypothetical prote
					cyclin 2b - Arabid

ALIGNMENTS

RESULT 1
BORT3
prostatic steroid-binding protein chain C3 precursor - rat
M:Alternate names: prostactin
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Dec-1982 #sequence_revision 14-Nov-1983 #text_change 28-May-1999
C:Accession: A92395; A92433; A91108; A42392; A03250
R:Paraker, M.G.; White, R.; Hurst, H.; Needham, M.; Tilly, R.
J. Biol. Chem. 258, 12-15, 1983
A:Title: Prostatic steroid-binding protein. Isolation and characterization of C3 gene
A:Reference number: A92395; MUID:83082848; PMID:6294095
A:Accession: A92395
A:Molecule type: mRNA
A:Residues: 1-95 <P>R</P>
A:Cross-references: GB:V01263; GB:J00777; MTD:g56993; PTD:CAA24577.1; PID:g56994
R:Viskochil, D. H.; Perry, S.T.; Lea, O.A.; Stafford, D.W.; Wilson, E.M.; French, F.S.
J. Biol. Chem. 258, 8861-8866, 1983
A:Title: Structural studies on rat prostatic binding protein. The primary structure o
A:Reference number: A91108; MUID:81188769; PMID:7014218
A:Accession: A91108
A:Molecule type: protein
A:Residues: 1-95 <P>E</P>
R:Pan, J.A.; Marschke, K.B.; Ho, K.C.; Perry, S.T.; Wilson, E.M.; French, F.S.
J. Biol. Chem. 267, 4456-4466, 1992
A:Title: Response elements of the androgen-regulated C3 gene.
A:Reference number: A42392; MUID:92165796; PMID:1537831
A:Accession: A42392
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-16, 'T', 18-95 <T>AN</T>
A:Note: sequence inconsistent with the nucleotide translation
A:Note: sequence extracted from NCBI backbone (NCBIN:83851, NCBI:83856)
C:Comment: C3 is encoded by two unique genes that differ from each other only in the
C:Comment: Steroid-binding protein, the principal androgen-dependent secretory protei
ng C2 and C3 chains. The chains of each dimer are linked by disulfide bonds. This pro
C:Superfamily: uteroglobin
C:Keywords: glycoprotein
F:1-18/Domain: signal sequence #status predicted <Sig>
F:19-95/Product: prostatic steroid-binding protein C3 chain #status experimental <MAT
F:35/Binding site: carbohydrate (Asn) (covalent) #status experimental
Query Match 39.4% Score 195.5; DB 1; Length 95;
Best local similarity 35.8%; Pred. No. 3.9e+13;
Matches 34; Conservative 30; Mismatches 30; Indels 1; Gaps 1;

QY 48 -----EFIDSDAAAEAMGKFKOCFLNOSHRTLNKNG 78
 Db 61 LDDYGIQVNEADADQADARVMAQVYDYGRCGSHPARLNLC 99

RESULT 6

cell division protein ftsy [imported] - Rickettsia conorii (strain/Malish 7)
 C:Species: Rickettsia conorii
 C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
 C:Accession: E97850
 R:Ogata, H.; Andic, S.; Renesto-Andiffren, P.; Fourrier, P.E.; Barbe, V.; Samson, D.; R.
 Science 293, 2093-2098, 2001
 A>Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893
 A:Accession: E97850
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-303 <KDP>
 A:Cross-references: GB:AF006914; PIDN:AAI03743.1; PID:g15620335; GSPDB:GM00173
 C:Genetics:
 A:Gene: ftsy
 C:Superfamily: docking protein

Query Match 14.3%; Score 71; DB 2; Length 303;
 Best Local Similarity 33.3%; Pred. No. 9.2;
 Matches 20; Conservative 13; Mismatches 17; Indels 10; Gaps 2;

QY 19 DSGCKLEDMVEKTI NSDISIPEYKELQEF-----IDSDAAEAMGKFKOCFLNOS 70
 Db 33 DAG--TLELELLISDSISVVTNIEEFKVKFKEDIDSTVKEALAKLEOOLSKS 90

RESULT 7

prostatic steroid-binding protein chain C2 precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 15-Oct-1982 #sequence_revision 15-Oct-1982 #text_change 24-Sep-1999
 C:Accession: A03251; A26671
 R:Parker, M.; Needham, M.; White, R.
 Nature 298, 92-94, 1982
 A>Title: Prostatic steroid binding protein: gene duplication and steroid binding.
 A:Reference number: A93286; MUID:82220075; PMID:6896562
 A:Accession: A03251
 A:Molecule type: mRNA
 A:Residues: 1-98 <PAR>
 A:Cross-references: GB:J00776; NID:q206448; PIDN:AA51641.1; PID:q206450
 R:Deleay, B.; Dirckx, L.; Decourt, J.L.; Claessens, F.; Peeters, B.; Rombaux, W.
 Nucleic Acids Res. 15, 1627-1641, 1987
 A>Title: Rat prostatic binding protein: the complete sequence of the C2 gene and its flanking regions.
 A:Reference number: A26671; MUID:87146484; PMID:2881277
 A:Accession: A26671

A:Molecule type: DNA
 A:Residues: 1-25, 'Q', 26-86, 'T', 88-94, 'VWLOINPGRWFSEIN'
 A:Cross-references: GB:X05034; NID:q56857; PIDN:CAA28708.1; PID:q56858
 C:Comment: Steroid-binding protein, the principal secretory protein in rat prostatic fluid.
 C:The chains of each dimer are linked by disulfide bonds.
 C:Superfamily: uteroglobin
 C:Keywords: heterotetramer; prostatic; steroid binding
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-98/Product: prostatic steroid-binding protein chain C2 #status predicted <MAP>

Query Match 14.1%; Score 70; DB 1; Length 98;
 Best Local Similarity 27.3%; Pred. No. 3.4;
 Matches 27; Conservative 20; Mismatches 36; Indels 16; Gaps 5;

QY 4 LNWLMALALHCHVADSG-----CKLEDM-VEKTI NSDISIPEYKELQEFIDSDAA 56
 Db 3 LSLCLITLIVCCYANQOTIAGYCOALQDVITIFLNP-EELEKLEEDAPPAEV 59

QY 57 EAMGKFKOCFLNOSHRTLNKFLMHT--VYDSIWMNK 93
 Db 60 EAMLKVKRCI-----NKIMYGDLSMGTSLVETMLKCDVK 94

RESULT 8
 B71638
 cell division protein ftsy (ftsY) RP775 - Rickettsia prowazekii

C:Species: Rickettsia prowazekii
 C>Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
 C:Accession: B71638
 R:Andersson, S.G.E.; Zomrodipour, A.; Andersson, J.O.; Sichteritz-Ponten, T.; Alsmark
 Nature 396, 133-140, 1998
 A>Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
 A:Reference number: A71630; MUID:99039499; PMID:9823893
 A:Accession: B71638
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-303 <AND>
 A:Cross-references: GB:AJ235273; GB:AJ235269; NID:q3861237; PIDN:CAI5202.1; PID:q386
 A:Experimental source: Strain Madrid E
 C:Genetics:
 A:Gene: ftsy; RP775
 C:Superfamily: docking protein

Query Match 14.1%; Score 70; DB 2; Length 303;
 Best Local Similarity 35.2%; Pred. No. 12;
 Matches 19; Conservative 10; Mismatches 17; Indels 8; Gaps 1;

QY 25 LEDMVEKTI NSDISIPEYKELQEF-----IDSDAAEAMGKFKOCFLNOS 70
 Db 37 LNFLEELISDSISVVTNIEEFKVKFKEDIDSTVKEALAKLEOOLSKS 90

RESULT 9

hypothetical protein T12K4.80 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T47321
 R:Montorf, A.; Casacuberta, E.; Pulidomenich, P.; Mewes, H.W.; Rudd, S.; Lemcke, K.;
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z24460
 A:Accession: T47321

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-412 <MON>
 A:Cross-references: EMBL:AL138640
 A:Experimental source: cultivar Columbia; BAC clone T12K4
 C:Genetics:
 A:Map position: 3
 A:Introns: 30/3
 A>Note: T12K4.80

Query Match 14.0%; Score 69.5; DB 2; Length 412;
 Best Local Similarity 27.5%; Pred. No. 18;
 Matches 25; Conservative 11; Mismatches 40; Indels 15; Gaps 3;

QY 13 LHCYVADSGCKLL--EDMVEKTI NSDISIPEYKELQEFID--SDAAEAMGKFKOCF 66
 Db 15 LLSCTYDSSAPIAKESKISRDSOKIQLVDYAPLVQITLSORRLPDVAHEIFLOTKSYN 74

QY 67 LNSHRTL-----KNFGIMHTVYDST 88
 Db 75 LIPNRTICALMLCFEANGFVLRARTWDEI 105

RESULT 10

protein PLIN9.15 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: C96669
 R:Theologis, A.; Becker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar,
 ansen, N.F.; Hughes, B.; Hultzer, L.

submitted to the Protein Sequence Database, January 2000

A:Reference number: 223037
 A:Accession: T46356
 A>Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-496 <AAA>
 A:Cross-references: EMBL:AL137289
 A:Experimental source: adult testis; clone DKFzp434G2016
 C:Genetics:
 A:Note: DKFzp434G2016.1

Query Match 13.8%; Score 68.5; DB 2; Length 496;
 Best Local Similarity 27.4%; Pred. No. 29;
 Matches 23; Conservative 15; Mismatches 29; Indels 17; Gaps 3;

OY 9 LAALLHCYADSGCKLLEDVVEKTN-SDISIPYKELL-----EEL-----OEFDIS 52
 Db 198 LTKLSKYVYFEACRLQKMDISDGFLLTFVQICKYPLQLAEILKTYHQHDFKDV 257

OY 53 DAAAEAKGKFKQCFNLQNSHRTLN 76
 Db 258 EAALHAMKNVAQ-LINERKRLRN 280

RESULT 15

H87369
 tryptophan halogenase, probable [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001

C:Accession: H87369

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.D.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MIMD:21173698; PMID:11259647

A:Accession: H87369

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-516 <STO>

A:Cross-references: GB:AE005673; NID:g13422252; PID:AAK22956.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0972

C:Superfamily: Rhizobium plasmid pMGR234a Y4XG protein

Query Match 13.7%; Score 68; DB 2; Length 516;
 Best Local Similarity 26.7%; Pred. No. 34;
 Matches 27; Conservative 18; Mismatches 40; Indels 16; Gaps 5;

OY 10 AALLHCYADSGC--KLLEDVVEKTN-SDISIPYKELL-----OEFDISDAAEAM 59
 Db 25 AAMLSHYQNGCAVELSESEIGTIGVSGSTIPPLQLLASLGVDEREFIOATQASPKL 84

OY 60 GKFKOCFLNOSHRTLNKNGILM-----MHTYDSIMCMKSN 95
 Db 85 GIRFEDMKQKQGRYYHPFGALGPIGHPHY-QCWLRARAKAN 124

Search completed: January 25, 2003, 13:35:21
 Job time : 60 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2003, 13:39:00 : Search time 20 seconds
(without alignments)
456.639 Million cell updates/sec

Title: US-09-806-302a-2

Perfect score: 95

Sequence: 1 MRLMVLMLALILHICVADS.....NFGIMHTVDSIWCNKS N 95

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : PIR-73:**

1: pir1:**

2: pir2:**

3: pir3:**

4: pir4:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9.5	705	2	S45769	probable membrane
2	8.4	249	2	T12463	hypothetical prote
3	8.4	256	2	H75404	outer membrane pro
4	8.4	552	1	E57987	cytochrome c-type
5	8.4	552	2	D86101	hypothetical prote
6	8.4	552	2	H91260	cytochrome c-type
7	8.4	937	2	P00864	hypothetical prote
8	8.4	2105	1	A44059	genome polypeptid
9	7.4	76	2	T03860	TA20 protein - com
10	7.4	107	2	E82494	conserved hypothet
11	7.4	239	2	A97358	glucose-inhibited
12	7.4	235	2	AB3613	glucose-1-phosphat
13	7.4	354	2	T18464	hypothetical prote
14	7.4	377	1	H64114	chorismate mutase
15	7.4	402	2	T15677	hypothetical prote
16	7.4	425	2	S58672	extra sex comb pr
17	7.4	433	2	S76553	hypothetical prote
18	7.4	572	2	T45139	chaperone protein
19	7.4	576	2	T36729	probable serine/th
20	7.4	605	2	G95853	probable pyruvate
21	7.4	627	2	H86254	hypothetical prote
22	7.4	665	2	F95053	cell wall surface
23	7.4	693	2	F95657	probable protein k
24	7.4	758	2	B83319	ATP-binding protei
25	7.4	775	2	D86261	hypothetical prote
26	7.4	994	2	A70776	probable glnh - My
27	6.3	46	2	T46224	hypothetical prote
28	6.3	49	2	A81605	hypothetical prote
29	6.3	57	2	D97724	hypothetical prote

30	6	6.3	67	2	F90882	hypothetical prote
31	6	6.3	67	2	A85736	hypothetical prote
32	6	6.3	67	2	D64894	hypothetical prote
33	6	6.3	68	2	B84265	hypothetical prote
34	6	6.3	75	2	A55419	dopamine D3 recept
35	6	6.3	79	2	B30924	hypothetical 9.8K
36	6	6.3	81	2	S71962	tirefoil factor p52
37	6	6.3	81	2	C86710	hypothetical prote
38	6	6.3	89	2	S08056	hypothetical prote
39	6	6.3	92	2	T09146	late-embryogenesis
40	6	6.3	95	1	B0RT3	prostatic steroid
41	6	6.3	95	2	G71630	integration host f
42	6	6.3	95	2	T09264	embryonic abundant
43	6	6.3	95	2	H97835	integration host f
44	6	6.3	102	2	D64363	ribosomal protein
45	6	6.3	103	2	E40899	hypothetical prote

ALIGNMENTS

RESULT 1
S45769
probable membrane protein YBL035c - yeast (Saccharomyces cerevisiae)
N/Alternate names: hypothetical protein YBL0414
C/Species: Saccharomyces cerevisiae
C/Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 21-Jul-2000
C/Accession: S45769; S46567
R/Goffeau, A.; Joniaux, J.-L.; Purnelle, B.; Skala, J.; de Wergifosse, P.; van Dyck, submitted to the Protein Sequence Database, August 1994
A/Reference number: S45745
A/Accession: S45769
A/Molecule type: DNA
A/Residues: 1-705 <GOR>
A/Cross-references: EMBL:Z35796; NID:9536044; PID:9536045; MIPS:YBL035c
A/Experimental source: strain S288C
R/Skala, J.; van Dyck, L.; Purnelle, B.; Goffeau, A.
Yeast 10(Suppl.A), S13-S24, 1994
A>Title: The sequence of an 8.8 kb segment on the left arm of chromosome II from Sacc and bacterial GTP cyclohydrolase II
A/Reference number: S46565; MUID:94378718; PMID:8091857
A/Accession: S46567
A/Molecule type: DNA
A/Residues: 1-705 <SKA>
A/Cross-references: EMBL:X74738; NID:9511140; PIDN:CAAS2761.1; PID:9511143
A/Experimental source: strain S288C
C/Genetics:
A/Gene: SGD:POL12
A/Cross-references: SGD:S0000131; MIPS:YBL035c
A/Map position: 2L
C/Keywords: transmembrane protein
F/428-444/Domain: transmembrane #status predicted <TMM>
Query Match 9.5%; Score 9; DB 2; Length 705;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 44 ELLOEFTDS 52
DB 411 ELLOEFTDS 419
RESULT 2
T12463
hypothetical protein DKFZp5641052.1 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999
C/Accession: T12463
R/Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A/Reference number: Z17522
A/Accession: T12463
A/Status: preliminary

A:Molecule type: mRNA
 A:Residues: 1-249 <DDE>
 A:Cross-references: EMBL:AL080063
 A:Experimental source: fetal brain; clone DKFZp5641052
 A:Genetics:
 A:Note: DKFZp5641052.1

Query Match
 Best Local Similarity 100.0%; Score 8; DB 2; Length 249;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 40 PEYKELQ 47
 |||||
 Db 136 PEYKELQ 143

RESULT 3

H75404
 Outer membrane protein - *Deinococcus radiodurans* (strain R1)
 C:Species: *Deinococcus radiodurans*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: H75404
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: H75404

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-256 <WHI>

A:Cross-references: GB:AE001982; GB:AE000513; NID:g6459109; PIDN:AAPI0931.1; PID:g645911

A:Experimental source: strain R1

A:Genetics:

A:Gene: DR1359

A:Map position: 1

C:Superfamily: Lipoprotein-28

Query Match
 Best Local Similarity 100.0%; Score 8; DB 2; Length 256;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 LMLAALL 14
 |||||
 Db 5 LMLAALL 12

RESULT 4

H57987

Cytochrome c-type biogenesis protein nrfE - *Escherichia coli* (strain K-12)

C:Species: *Escherichia coli*

C:Date: 26-Jul-1996 #sequence_revision 31-Oct-1997 #text_change 01-Mar-2002

C:Accession: A65216; E57987

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
 a.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.

A:Reference number: A64720; MUID:97426617; PMID:9276503

A:Accession: A65216

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-552 <BLAT>

A:Cross-references: GB:AE000481; GB:U00096; NID:g2367346; PIDN:AMD14457.1; PID:g1790511;

A:Experimental source: strain K-12, substrain MG1655

R:Husain, H.; Grove, J.; Griffiths, L.; Busby, S.; Cole, J.

Mol. Microbiol. 12, 153-163, 1994

A:Title: A seven gene operon essential for formate-dependent nitrite reduction to ammoni

A:Reference number: A57987; MUID:94355626; PMID:8057835

A:Accession: E57987

A:Molecule type: DNA

A:Residues: 1-31, 'V', '33-37', 'A', '39-189', 'S', '191-217', 'L', '219', 'RLVGLS', '227-454', 'A', '456-480',
 A:Cross-references: EMBL:X72298; NID:g404302; PIDN:CAA51045.1; PID:g581146

C:Genetics:
 A:Gene: nrfE
 A:Note: this is one of the seven gene operon (nrf) essential for formate-dependent n
 C:Function:
 A:Description: an inner membrane protein, possible component of a heme lyase, which i
 ent nitrite reductase activity
 C:Superfamily: nrfE protein
 C:Keywords: membrane protein

Query Match
 Best Local Similarity 100.0%; Score 8; DB 1; Length 552;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 LMLAALL 14
 |||||
 Db 114 LMLAALL 121

RESULT 5

H6101

Hypothetical protein nrfE [imported] - *Escherichia coli* (strain O157:H7, substrain ED

C:Species: *Escherichia coli*

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Dec-2001

C:Accession: H6101

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Ma
 y
 11ter, L.; Grobeck, E.J.; Davis, N.W.; Lin, A.; Dimaienta, E.; Potamouisis, K.; Apoda
 Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: H6101

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-552 <SPD>

A:Cross-references: GB:AE005174; NID:g12519031; PIDN:AAC59272.1; GSPDB:GN00145; UNCP:

A:Experimental source: strain O157:H7, substrain EDU933

A:Genetics:

A:Gene: nrfE

C:Superfamily: nrfE protein

Query Match
 Best Local Similarity 100.0%; Score 8; DB 2; Length 552;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 LMLAALL 14
 |||||
 Db 114 LMLAALL 121

RESULT 6

H91260

Cytochrome c-type biogenesis protein nrfE [imported] - *Escherichia coli* (strain O157:

C:Species: *Escherichia coli*

C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 14-Dec-2001

C:Accession: H91260

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinozawa, H.
 DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and g

A:Reference number: A96629; MUID:21156231; PMID:11258756

A:Accession: H91260

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-552 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA838479.1; PID:g13364533; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RMD 0509952

C:Genetics:

A:Gene: EC85056

C:Superfamily: nrfE protein

Query Match
 Best Local Similarity 100.0%; Score 8; DB 2; Length 552;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LMALALL 14
 DB 114 LMALALL 121

RESULT 7
 PQ0864
 hypothetical protein 937 - citrus tatter leaf virus (fragment)

N:Alternate names: ORF1
 C:Species: citrus tatter leaf virus
 C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 20-Jun-2000

C:Accession: PQ0864
 R:Yoshikawa, N.; Imaizumi, M.; Takahashi, T.; Inouye, N.
 J. Gen. Virol. 74, 2743-2747, 1993

A:Title: Striking similarities between the nucleotide sequence and genome organization of
 A:Reference number: PQ0864; MWID:94103780; PMID:1277280

A:Accession: PQ0864

A:Molecule type: mRNA

A:Residues: 1-937 <YOS>

A:Cross-references: DDBJ:D16368; NID:g464150; PIDN:BA03869.1; PID:g563955

A:Experimental source: strain L1-23

C:Superfamily: eggplant mosaic virus RNA-directed RNA polymerase

C:Keywords: ATP

Query Match 8.4% Score 8; DB 2; Length 937;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 KELLOEFT 50
 DB 55 KELLOEFT 62

RESULT 8

AA4059

genome polyprotein - apple stem grooving virus (strain P-209)

N:Alternate names: RNA nucleotidyltransferase (RNA-directed); RNA replicase

N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)

C:Species: apple stem grooving virus, ASGV

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 19-Jan-2001

C:Accession: AA4059

R:Yoshikawa, N.; Sasaki, E.; Kato, M.; Takahashi, T.

Virol. 191, 98-105, 1992

A:Title: The nucleotide sequence of apple stem grooving capilliovirus genome.

A:Reference number: AA4059; MWID:93033164; PMID:1413530

A:Accession: AA4059

A:Molecule type: genomic RNA

A:Residues: 1-2105 <YOS>

A:Cross-references: GB:D14995; NID:g303496; PIDN:BA03639.1; PID:g285608

C:Superfamily: eggplant mosaic virus RNA-directed RNA polymerase

C:Keywords: ATP; nucleotide binding; nucleotidyltransferase; P-loop; RNA biosynthesis; F

F:781-188/Region: nucleotide-binding motif A (P-loop)

F:843-148/Region: nucleotide-binding motif B

F:787/Binding site: ATP (Lys) #status predicted

Query Match 8.4% Score 8; DB 1; Length 2105;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 KELLOEFT 50
 DB 1223 KELLOEFT 1230

RESULT 9

T03860

TA20 protein - common tobacco

C:Species: Nicotiana tabacum (common tobacco)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 08-Oct-1999

C:Accession: T03860

R:Beals, T. P.; Goldberg, R. B.

submitted to the EMBL Data Library, October 1996

A:Description: Nicotiana tabacum gene expressed in anther.

A:Reference number: 215122

A:Accession: T03860

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-76 <BEA>

A:Cross-references: EMBL:U73164; NID:g1657813; PIDN:AA818190.1; PID:g1657814

A:Experimental source: tissue-type anther

C:Genetics:

A:Gene: TA20

Query Match 7.4% Score 7; DB 2; Length 76;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LMVLMIA 10
 DB 19 LMVLMIA 25

RESULT 10

E82494

conserved hypothetical protein VCA0152 [imported] - Vibrio cholerae (strain N16961 se

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 16-Feb-2001

C:Accession: E82494

R:Heideberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.

Chardom, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MWID:20406833; PMID:10952301

A:Accession: E82494

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-107 <HEI>

A:Cross-references: GB:AE004356; GB:AE003853; NID:g9657536; PIDN:AAF96065.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VCA0152

A:Map position: 2

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1428

Query Match 7.4% Score 7; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 MLALALL 14
 DB 3 MLALALL 9

RESULT 11

A97358

glucose-inhibited division protein, GIDB (S-adenosylmethionine-dependent methyltransferase

C:Species: Clostridium acetobutylicum

C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: A97358

R:Kolling, J.; Breton, G.; Omejenko, M.V.; Markkova, K.S.; Zeng, Q.; Gibson, R.; L.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium

A:Reference number: A96900; MWID:21359325; PMID:21359325

A:Accession: A97358

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-219 <KOR>

A:Cross-references: GB:AE001437; PIDN:AAK81652.1; PID:g15026839; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC3732

Query Match 7.4% Score 7; DB 2; Length 239;
 Best Local Similarity 100.0%; Pred. No. 22;

Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 YKELQE 48
|||||||
Db 29 YKELQE 35

RESULT 12

AB3613
glucose-1-phosphate cytidylyltransferase (EC 2.7.7.33) [Imported] - Brucella melitensis
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C:Accession: AB3613
R:DelVecchio, V.G.; Kapurali, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AB3613
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <KUR>
A:Cross-references: GB:AE008918; PIDN:AAL54069.1; PID:g17985025; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0827
A:Map position: 11
C:Superfamily: glucose-1-phosphate cytidylyltransferase
C:Keywords: nucleocytoplasmic transferase

Query Match 7.4%; Score 7; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 LEDWEK 31
|||||||
Db 268 LEDWEK 274

RESULT 13

T18464
hypothetical protein C0480C - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2002
C:Accession: T18464
R:Lawson, D.; Bowman, S.; Bartell, B.
Submitted to the EMBL Data Library, November 1998
A:Reference number: Z18937
A:Accession: T18464
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-354 <LAW>
A:Cross-references: EMBL:AL008970; NID:e1407852; PIDN:CAI5607.1
C:Genetics:
A:Map position: 3
A:Note: C0480C

Query Match 7.4%; Score 7; DB 2; Length 354;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 LQEFIDS 52
|||||||
Db 268 LQEFIDS 274

RESULT 14

H64114
chorismate mutase (EC 5.4.99.5) / prephenate dehydrogenase (EC 1.3.1.12) H11290 [similar
N:Contains: chorismate mutase T; prephenate dehydrogenase
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 10-May-1996 #text_change 06-Oct-2000
C:Accession: H64114; T09419

R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.T.; Glodek, A.; Kelley, J.M.; McDaniel,
D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fritchman, J.L.; Geoghegan, N.S.M.
Science 269, 496-512, 1995
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A:Reference number: A64000; M01D:95350630; PMID:7542800
A:Accession: H64114
A:Molecule type: DNA

A:Residues: 1-377 <RTGR>
A:Cross-references: GB:U32809; GB:L42023; NID:g3212217; PIDN:MAC22939.1; PID:g1574749
A:Note: named as homolog to a protein from Erwinia herbicola
R:White, O.; Clayton, R.A.; Kerlavage, A.R.; Fleischmann, R.D.; Peterson, J.; Hickey,
submitted to the EMBL Data Library, May 1998
A:Reference number: Z16667
A:Accession: T09419
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-377 <WHI>
A:Cross-references: EMBL:U32809; NID:g3212217; PIDN:MAC22939.1; PID:g1574749
C:Genetics:
A:Gene: H11290
A:Start codon: GTG
C:Superfamily: tyra bifunctional enzyme
C:Keywords: intramolecular transferase; isomerase; multifunctional enzyme; oxidoreduc

Query Match 7.4%; Score 7; DB 1; Length 377;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 KTINSDI 37
|||||||
Db 94 KTINSDI 100

RESULT 15

T15677
hypothetical protein C28C12.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15677
R:Miller, N.
Submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C28C12.
A:Reference number: Z18387
A:Accession: T15677
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-402 <MTL>
A:Cross-references: EMBL:U40797; NID:g1065916; PID:g1065921; PIDN:AAB37548.1; GSPDB:G
A:Experimental source: strain Bristol N2; clone C28C12
C:Genetics:
A:Gene: CESP:C28C12.5
A:Map position: 4
A:Introns: 29/3; 82/3; 124/3; 151/3; 258/3; 318/1; 372/3; 400/3

Query Match 7.4%; Score 7; DB 2; Length 402;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 LLOEFTD 51
|||||||
Db 250 LLOEFTD 256

Search completed: January 25, 2003, 13:41:50
Job time : 23 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2003, 13:34:15 : Search time 10 Seconds
(without alignments)
191.696 Million cell updates/sec

Title: US-09-806-302a-2

Perfect score: 95

Sequence: 1 MKLMLVLAALILHCYADS.....NFGMLMHTVYDSIMCNMKN 95

Scoring table: ORIGO

Gapop 60.0 , Capext 60.0

Searched: 122226 seqs, 20178551 residues

Word size : 0

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCU05_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/PCU05_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	95	10	US-09-110-716-31
2	95	100.0	95	10	US-09-985-911-6
3	77	81.1	77	10	US-09-110-716-13
4	65	68.4	76	10	US-09-110-716-40
5	12	12.6	13	10	US-09-757-417-29
6	12	12.6	93	10	US-09-757-417-27
7	12	12.6	93	10	US-09-934-054-3
8	12	12.6	93	10	US-09-934-054-10
9	12	12.6	93	12	US-10-007-805-503
10	12	12.6	132	10	US-09-757-417-47
11	12	12.6	410	12	US-10-007-805-495
12	12	12.6	743	12	US-10-007-805-484
13	12	12.6	1095	12	US-10-007-805-493
14	11	11.6	20	10	US-09-757-417-4
15	11	11.6	20	10	US-09-757-417-33
16	11	11.6	20	12	US-10-007-805-499
17	11	11.6	75	10	US-09-110-716-41
18	10	10.5	10	10	US-09-757-417-42
19	10	10.5	20	10	US-09-757-417-32

20	9	9.5	9	10	US-09-757-417-5	Sequence 5, Appl1
21	9	9.5	9	10	US-09-757-417-37	Sequence 37, Appl
22	9	9.5	9	10	US-09-757-417-38	Sequence 38, Appl
23	9	9.5	9	10	US-09-757-417-39	Sequence 39, Appl
24	9	9.5	9	12	US-10-007-805-500	Sequence 500, App
25	8	8.4	20	10	US-09-757-417-12	Sequence 12, Appl
26	8	8.4	20	10	US-09-757-417-14	Sequence 14, Appl
27	8	8.4	20	10	US-09-757-417-15	Sequence 15, Appl
28	8	8.4	20	10	US-09-757-417-17	Sequence 17, Appl
29	7	7.4	10	10	US-09-757-417-45	Sequence 45, Appl
30	7	7.4	10	10	US-09-864-761-45420	Sequence 45420, A
31	7	7.4	112	10	US-09-682-706-2	Sequence 2, Appl1
32	7	7.4	112	10	US-09-919-473-2	Sequence 2, Appl1
33	6	6.3	9	10	US-09-244-694-184	Sequence 184, App
34	6	6.3	22	9	US-09-865-989-11	Sequence 11, Appl
35	6	6.3	22	9	US-09-865-989-17	Sequence 17, Appl
36	6	6.3	22	9	US-09-865-989-23	Sequence 23, Appl
37	6	6.3	22	9	US-09-865-989-36	Sequence 36, Appl
38	6	6.3	22	9	US-09-865-989-87	Sequence 87, Appl
39	6	6.3	31	10	US-09-732-091-13	Sequence 13, Appl
40	6	6.3	40	10	US-09-925-299-1468	Sequence 1468, Ap
41	6	6.3	48	10	US-09-864-761-40519	Sequence 40519, A
42	6	6.3	63	10	US-09-864-761-34909	Sequence 34909, A
43	6	6.3	70	10	US-09-289-346A-1	Sequence 1, Appl1
44	6	6.3	70	10	US-09-289-346A-2	Sequence 2, Appl1
45	6	6.3	70	10	US-09-289-346A-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-110-716-31
: Sequence 31, Application US/09110716A
: Patent No. US2002034739A1
: GENERAL INFORMATION:
: APPLICANT: Lehrer, Robert I.
: APPLICANT: Zhao, Chengquan
: APPLICANT: Glasgow, Benjamin J.
: TITLE OF INVENTION: PEPTIDES CHARACTERISTIC OF CERTAIN TUMORS
: FILE REFERENCE: 22000-20596.00
: CURRENT APPLICATION NUMBER: US/09/110.716A
: NUMBER OF SEQ ID NOS: 41
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 31
: LENGTH: 95
: TYPE: PRT
: ORGANISM: lipophilin C
US-09-110-716-31
Query Match 100.0%; Score 95; DB 10; Length 95;
Best Local Similarity 100.0%; Pred. No. 3.2e-83;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MKLMLVLAALILHCYADSGCKLLEDEYVEKTIISDIPEYKELLOERIDSDAAAEAMG 60
DB 1 MKLMLVLAALILHCYADSGCKLLEDEYVEKTIISDIPEYKELLOERIDSDAAAEAMG 60
OY 61 KFKOCFLNOSHRTLNKFGIMHFTVYDSIMCNMKN 95
DB 61 KFKOCFLNOSHRTLNKFGIMHFTVYDSIMCNMKN 95
RESULT 2
US-09-985-911-6
: Sequence 6, Application US/09985911
: Patent No. US20020151012A1
: GENERAL INFORMATION:
: APPLICANT: NI ET AL.
: TITLE OF INVENTION: HUMAN ENDOMETRIAL SPECIFIC STEROID-BINDING FACTOR I, II AND III
: FILE REFERENCE: PF57573
: CURRENT APPLICATION NUMBER: US/09/985.911

```

; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: 09/583,169
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: 09/263,810
; PRIOR FILING DATE: 1999-03-08
; PRIOR APPLICATION NUMBER: 08/821,451
; PRIOR FILING DATE: 1997-03-21
; PRIOR APPLICATION NUMBER: 60/014,724
; PRIOR FILING DATE: 1996-03-21*
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 95
; TYPE: PRT
; ORGANISM: human
US-09-985-911-6

Query Match
Best Local Similarity 100.0%; Score 95; DB 10; Length 95;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLWMLAALLHCYADSCGKLEDMVEKTI NSDISIPEYKELLQEFIDSDAAAEAMG 60
DB 1 MKLWMLAALLHCYADSCGKLEDMVEKTI NSDISIPEYKELLQEFIDSDAAAEAMG 60
QY 61 KFKOCFLNOSHRTLNKFLMHTYDSTWCMKSN 95
DB 61 KFKOCFLNOSHRTLNKFLMHTYDSTWCMKSN 95

RESULT 3
US-09-110-716-13
; Sequence 13, Application US/09110716A
; Patent No. US20020034739A1
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Zhao, Chengquan
; APPLICANT: Glasgow, Benjamin J.
; TITLE OF INVENTION: PEPTIDES CHARACTERISTIC OF CERTAIN TUMORS
; FILE REFERENCE: 22000-20596.00
; CURRENT APPLICATION NUMBER: US/09/110,716A
; CURRENT FILING DATE: 1998-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-110-716-13

Query Match
Best Local Similarity 81.1%; Score 77; DB 10; Length 77;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 DSGCKLEDMVEKTI NSDISIPEYKELLQEFIDSDAAAEAMGKFCFLNOSHRTLNKNG 78
DB 1 DSGCKLEDMVEKTI NSDISIPEYKELLQEFIDSDAAAEAMGKFCFLNOSHRTLNKNG 60
QY 79 LMMHTYDSTWCMKSN 95
DB 61 LMMHTYDSTWCMKSN 77

RESULT 4
US-09-110-716-40
; Sequence 40, Application US/09110716A
; Patent No. US20020034739A1
; GENERAL INFORMATION:
; APPLICANT: Lehrer, Robert I.
; APPLICANT: Zhao, Chengquan
; APPLICANT: Glasgow, Benjamin J.
; TITLE OF INVENTION: PEPTIDES CHARACTERISTIC OF CERTAIN TUMORS
; FILE REFERENCE: 22000-20596.00
```

```

; CURRENT APPLICATION NUMBER: US/09/110,716A
; CURRENT FILING DATE: 1998-07-07
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 40
; LENGTH: 76
; TYPE: PRT
; ORGANISM: LPNC
US-09-110-716-40

Query Match
Best Local Similarity 68.4%; Score 65; DB 10; Length 76;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 DSGCKLEDMVEKTI NSDISIPEYKELLQEFIDSDAAAEAMGKFCFLNOSHRTLNKNG 78
DB 1 DSGCKLEDMVEKTI NSDISIPEYKELLQEFIDSDAAAEAMGKFCFLNOSHRTLNKNG 60
QY 79 LMMHT 83
DB 61 LMMHT 65

RESULT 5
US-09-757-417-29
; Sequence 29, Application US/09757417
; Patent No. US20020082216A1
; GENERAL INFORMATION:
; APPLICANT: Fanger, Gary R.
; APPLICANT: Foy, Theresa M.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY, DIAGNOSIS AND MONITORING OF BREAST CANCER
; FILE REFERENCE: 210121.479C1
; CURRENT APPLICATION NUMBER: US/09/757,417
; CURRENT FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-417-29

Query Match
Best Local Similarity 12.6%; Score 12; DB 10; Length 13;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLWMLAAL 12
DB 1 MKLWMLAAL 12

RESULT 6
US-09-757-417-27
; Sequence 27, Application US/09757417
; Patent No. US20020082216A1
; GENERAL INFORMATION:
; APPLICANT: Fanger, Gary R.
; APPLICANT: Foy, Theresa M.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY, DIAGNOSIS AND MONITORING OF BREAST CANCER
; FILE REFERENCE: 210121.479C1
; CURRENT APPLICATION NUMBER: US/09/757,417
; CURRENT FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 93
; TYPE: PRT
```

```

: ORGANISM: Homo sapien
US-09-757-417-27

Query Match      12.6%; Score 12; DB 10; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKLIMVIMLAL 12
      |||
Db      1 MKLIMVIMLAL 12

RESULT 7
US-09-934-054-3
: Sequence 3, Application US/09934054
: Patent No. US20020107385A1
: GENERAL INFORMATION:
: APPLICANT: Aetblom, Ingrid E.
: Hillman, Jennifer L.
: Moriy, Lynn E.
: Goll, Surya K.
: Hawkins, Phillip R.
: TITLE OF INVENTION: BREAST TUMOR SPECIFIC PROTEINS
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 2200 Sand Hill Road, Suite 100
: CITY: Menlo Park
: STATE: CA
: COUNTRY: USA
: ZIP: 94025-6936
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/934,054
: FILING DATE: 21-Aug-2001
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/747,547
: FILING DATE: <unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0077 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 855-0555
: TELEFAX: (650) 845-4166
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 93 amino acids
: TYPE: amino acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: <unknown>
: SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-934-054-3

Query Match      12.6%; Score 12; DB 10; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKLIMVIMLAL 12
      |||
Db      1 MKLIMVIMLAL 12

RESULT 8
US-09-934-054-10
: Sequence 10, Application US/09934054
: Patent No. US20020107385A1
: GENERAL INFORMATION:
: APPLICANT: Aetblom, Ingrid E.
: Hillman, Jennifer L.
: Moriy, Lynn E.
: Goll, Surya K.
: Hawkins, Phillip R.
: TITLE OF INVENTION: BREAST TUMOR SPECIFIC PROTEINS
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 2200 Sand Hill Road, Suite 100
: CITY: Menlo Park
: STATE: CA
: COUNTRY: USA
: ZIP: 94025-6936
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/934,054
: FILING DATE: 21-Aug-2001
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/747,547
: FILING DATE: <unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0077 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 855-0555
: TELEFAX: (650) 845-4166
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 93 amino acids
: TYPE: amino acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: <unknown>
: SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-934-054-10

Query Match      12.6%; Score 12; DB 10; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKLIMVIMLAL 12
      |||
Db      1 MKLIMVIMLAL 12

RESULT 9
US-10-007-805-503
: Sequence 503, Application US/10007805
: Patent No. US20020150581A1
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugu
: APPLICANT: Dillon, Davin C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Xu, Jlangchun
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Hepler, William T.
: APPLICANT: Henderson, Robert A.
: APPLICANT: Fanger, Gary R.
: APPLICANT: Vedrick, Thomas S.
: APPLICANT: McNeill, Patricia D.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
: FILE REFERENCE: 210121.470C10
: CURRENT APPLICATION NUMBER: US/10/007,805
: CURRENT FILING DATE: 2001-12-07
: NUMBER OF SEQ ID NOS: 593
```

```

: APPLICANT: Aetblom, Ingrid E.
: Hillman, Jennifer L.
: Moriy, Lynn E.
: Goll, Surya K.
: Hawkins, Phillip R.
: TITLE OF INVENTION: BREAST TUMOR SPECIFIC PROTEINS
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 2200 Sand Hill Road, Suite 100
: CITY: Menlo Park
: STATE: CA
: COUNTRY: USA
: ZIP: 94025-6936
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/934,054
: FILING DATE: 21-Aug-2001
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/747,547
: FILING DATE: <unknown>
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0077 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 855-0555
: TELEFAX: (650) 845-4166
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 93 amino acids
: TYPE: amino acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: <unknown>
: SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-934-054-10

Query Match      12.6%; Score 12; DB 10; Length 93;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MKLIMVIMLAL 12
      |||
Db      1 MKLIMVIMLAL 12

RESULT 9
US-10-007-805-503
: Sequence 503, Application US/10007805
: Patent No. US20020150581A1
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugu
: APPLICANT: Dillon, Davin C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Xu, Jlangchun
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Hepler, William T.
: APPLICANT: Henderson, Robert A.
: APPLICANT: Fanger, Gary R.
: APPLICANT: Vedrick, Thomas S.
: APPLICANT: McNeill, Patricia D.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
: FILE REFERENCE: 210121.470C10
: CURRENT APPLICATION NUMBER: US/10/007,805
: CURRENT FILING DATE: 2001-12-07
: NUMBER OF SEQ ID NOS: 593
```

```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 503
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-503
```

```
Query Match
Best Local Similarity 100.0%; Score 12; DB 12; Length 93;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MKLWVLMAL 12
Db 1 MKLWVLMAL 12
```

```
RESULT 10
US-09-757-417-47
; Sequence 47, Application US/09757417
; Patent No. US20020082216A1
; GENERAL INFORMATION:
; APPLICANT: Fanger, Gary R.
; APPLICANT: Foy, Theresa M.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY, DIAGNOSIS AND MONITORING OF BREAST CANCER
; FILE REFERENCE: 210121.479C1
; CURRENT APPLICATION NUMBER: US/09/757,417
; FILE REFERENCE: 210121.479C1
; CURRENT FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 132
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-757-417-47
```

```
Query Match
Best Local Similarity 100.0%; Score 12; DB 10; Length 132;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MKLWVLMAL 12
Db 40 MKLWVLMAL 51
```

```
RESULT 11
US-10-007-805-495
; Sequence 495, Application US/10007805
; Patent No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaogchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 495
; LENGTH: 410
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
US-10-007-805-495
```

```
Query Match
Best Local Similarity 100.0%; Score 12; DB 12; Length 410;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MKLWVLMAL 12
Db 1 MKLWVLMAL 12
```

```
RESULT 12
US-10-007-805-494
; Sequence 494, Application US/10007805
; Patent No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaogchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 494
; LENGTH: 743
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 403..522, 615
; OTHER INFORMATION: Xaa - Any Amino Acid
US-10-007-805-494
```

```
Query Match
Best Local Similarity 100.0%; Score 12; DB 12; Length 743;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 MKLWVLMAL 12
Db 1 MKLWVLMAL 12
```

```
RESULT 13
US-10-007-805-493
; Sequence 493, Application US/10007805
; Patent No. US20020150581A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaogchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margarita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C10
```


THIS PAGE BLANK (0870)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2003, 13:27:44 ; Search time 18 Seconds
(without alignments)
135,288 Million cell updates/sec

Title: US-09-806-302a-2

Perfect score: 95

Sequence: 1 MKLIMVLMALLHCYADSDS.....NFGIMHTVDSIMCNKSN 95

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/laa/5a.COMB.pep:*
- 2: /cgn2_6/ptodata/2/laa/5b.COMB.pep:*
- 3: /cgn2_6/ptodata/2/laa/5c.COMB.pep:*
- 4: /cgn2_6/ptodata/2/laa/5d.COMB.pep:*
- 5: /cgn2_6/ptodata/2/laa/5e.COMB.pep:*
- 6: /cgn2_6/ptodata/2/laa/5f.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	95	100.0	95	3	US-08-821-451A-6
2	95	100.0	95	4	US-09-263-810-6
3	95	100.0	95	4	US-09-583-169-6
4	12	12.6	93	1	US-08-455-896-2
5	12	12.6	93	2	US-08-933-149-2
6	12	12.6	93	2	US-09-082-343-2
7	12	12.6	93	3	US-09-082-253-2
8	12	12.6	93	4	US-09-215-818-5
9	12	12.6	93	5	PCT-US96-08235-2
10	11	7.4	528	2	US-08-403-852D-21
11	7	7.4	528	3	US-08-510-646B-22
12	7	7.4	528	4	US-09-231-818-21
13	6	6.3	11	1	US-07-699-468-2
14	6	6.3	11	1	US-07-699-468-3
15	6	6.3	11	1	US-07-699-468-4
16	6	6.3	11	1	US-07-699-468-5
17	6	6.3	11	1	US-07-699-468-6
18	6	6.3	11	2	US-08-723-415B-9
19	6	6.3	16	4	US-09-189-627A-9
20	6	6.3	16	4	US-09-710-861-9
21	6	6.3	22	3	US-08-940-095-11
22	6	6.3	22	3	US-08-940-095-17
23	6	6.3	22	3	US-08-940-095-23
24	6	6.3	22	3	US-08-940-095-36
25	6	6.3	22	3	US-08-940-095-87
26	6	6.3	22	3	US-08-940-093-11
27	6	6.3	22	3	US-08-940-093-17

28	6	6.3	22	3	US-08-940-093-23	Sequence 23, Appl
29	6	6.3	22	3	US-08-940-093-36	Sequence 36, Appl
30	6	6.3	22	3	US-08-940-093-87	Sequence 87, Appl
31	6	6.3	22	3	US-08-940-096-11	Sequence 11, Appl
32	6	6.3	22	3	US-08-940-096-17	Sequence 17, Appl
33	6	6.3	22	3	US-08-940-096-23	Sequence 23, Appl
34	6	6.3	22	3	US-08-940-096-36	Sequence 36, Appl
35	6	6.3	22	3	US-08-940-096-87	Sequence 87, Appl
36	6	6.3	22	4	US-09-465-719-11	Sequence 11, Appl
37	6	6.3	22	4	US-09-465-719-17	Sequence 17, Appl
38	6	6.3	22	4	US-09-465-719-23	Sequence 23, Appl
39	6	6.3	22	4	US-09-465-719-36	Sequence 36, Appl
40	6	6.3	22	4	US-09-465-719-87	Sequence 87, Appl
41	6	6.3	22	4	US-09-453-605-11	Sequence 11, Appl
42	6	6.3	22	4	US-09-453-605-17	Sequence 17, Appl
43	6	6.3	22	4	US-09-453-605-23	Sequence 23, Appl
44	6	6.3	22	4	US-09-453-605-36	Sequence 36, Appl
45	6	6.3	22	4	US-09-453-605-87	Sequence 87, Appl

ALIGNMENTS

```

RESULT 1
US-08-821-451A-6
; sequence 6, Application US/08821451A
; Patent No. 6066724
; GENERAL INFORMATION:
; APPLICANT: Jian NI, Guo-Liang Yu and Reiner Gentz
; TITLE OF INVENTION: Human Endometrial Specific Steroid-
; TITLE OF INVENTION: Binding Factor I, II and III
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/821,451A
; FILING DATE: March 21, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,724
; FILING DATE: March 21, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.C.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-521 (PP257)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDNESS:
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-08-821-451A-6
;
Query Match 100.0%; Score 95; DB 3; Length 95;
Best Local Similarity 100.0%; Pred. No. 2.8e-87;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MKLIMVLMALLHCYADSDCKLLEDMVEKTTNSDISIPYKELLQERTDSDAALAEKMG 60

```

DB 1 MKLWVLMALALLHCHYADSCKLEDMVEKTIINSISIPREKELLOEFTDSDAAEAMG 60
QY 61 KFKOCFLNOSHRTLKNFGLMHMTYVDSIMCNMKS 95
DB 61 KFKOCFLNOSHRTLKNFGLMHMTYVDSIMCNMKS 95

RESULT 2

US-09-263-810-6
Sequence 6, Application US/09263810
Patent No. 6174992
GENERAL INFORMATION:
APPLICANT: JIAN N1, Guo-Liang Yu and Reiner Gentz
TITLE OF INVENTION: Human Endometrial Specific Steroid-
TITLE OF INVENTION: Binding Factor I, II and III
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/263,810
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/821,451
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-521 (PF257)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-09-263-810-6

Query Match 100.0%; Score 95; DB 4; Length 95;
Best Local Similarity 100.0%; Pred. No. 2.8e-87;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLWVLMALALLHCHYADSCKLEDMVEKTIINSISIPREKELLOEFTDSDAAEAMG 60
DB 1 MKLWVLMALALLHCHYADSCKLEDMVEKTIINSISIPREKELLOEFTDSDAAEAMG 60
QY 61 KFKOCFLNOSHRTLKNFGLMHMTYVDSIMCNMKS 95
DB 61 KFKOCFLNOSHRTLKNFGLMHMTYVDSIMCNMKS 95

RESULT 3

US-09-583-169-6
Sequence 6, Application US/09583169
Patent No. 6338948
GENERAL INFORMATION:
APPLICANT: JIAN N1, Guo-Liang Yu and Reiner Gentz
TITLE OF INVENTION: Human Endometrial Specific Steroid-

TITLE OF INVENTION: Binding Factor I, II and III
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/583,169
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/821,451
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-521 (PF257)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-09-583-169-6

Query Match 100.0%; Score 95; DB 4; Length 95;
Best Local Similarity 100.0%; Pred. No. 2.8e-87;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLWVLMALALLHCHYADSCKLEDMVEKTIINSISIPREKELLOEFTDSDAAEAMG 60
DB 1 MKLWVLMALALLHCHYADSCKLEDMVEKTIINSISIPREKELLOEFTDSDAAEAMG 60
QY 61 KFKOCFLNOSHRTLKNFGLMHMTYVDSIMCNMKS 95
DB 61 KFKOCFLNOSHRTLKNFGLMHMTYVDSIMCNMKS 95

RESULT 4

US-08-455-896-2
Sequence 2, Application US/08455896
Patent No. 5668267
GENERAL INFORMATION:
APPLICANT: WATSON, MARK A.
APPLICANT: FLEMING, TIMOTHY P.
TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
TITLE OF INVENTION: NANOMARY-SPECIFIC BREAST CANCER PROTEIN
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROGERS, HOWELL & HAPERKAMP
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: USA
ZIP: 63105-1817
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,896
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 952726
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-455-896-2

Query Match 12.6%; Score 12; DB 1; Length 93;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLWVLMIAL 12
Db 1 MKLWVLMIAL 12

RESULT 5

US-08-933-149-2
Sequence 2, Application US/08933149
Patent No. 5922836

GENERAL INFORMATION:
APPLICANT: WATSON, MARK A.
APPLICANT: FLEMING, TIMOTHY P.
TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN
TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: USA
ZIP: 63105-1817

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/933,149
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: HENDERSON, MELDIE W.
REGISTRATION NUMBER: 37,846
REFERENCE/DOCKET NUMBER: 6029-6040
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-08-933-149-2

Query Match 12.6%; Score 12; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLWVLMIAL 12
Db 1 MKLWVLMIAL 12

RESULT 6

US-09-082-343-2
Sequence 2, Application US/09082343
Patent No. 5968754

GENERAL INFORMATION:
APPLICANT: WATSON, MARK A.
APPLICANT: FLEMING, TIMOTHY P.
TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROGERS, HOWELL & HAFERKAMP
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: USA
ZIP: 63105-1817
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,343
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/455,896
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 952726
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-082-343-2

Query Match 12.6%; Score 12; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKLWVLMIAL 12
Db 1 MKLWVLMIAL 12

RESULT 7

US-09-082-253-2
Sequence 2, Application US/09082253
Patent No. 6004756

GENERAL INFORMATION:
APPLICANT: WATSON, MARK A.
APPLICANT: FLEMING, TIMOTHY P.
TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN

NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROGERS, HOWELL & HAFERKAMP
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: USA
ZIP: 63105-1817
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,253
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/455,896
FILING DATE: 05/31/1995
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 952726
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
US-09-082-253-2

Query Match 12.6%; Score 12; DB 3; Length 93;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLWVLMAL 12
DB 1 MKLWVLMAL 12

RESULT 8
US-09-215-818-5
Sequence 5, Application US/09/215,818A
Patent No. 6379671
GENERAL INFORMATION:
APPLICANT: Colpits, Tracey
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL FOR
FILE OF INVENTION: DETECTING DISEASES OF THE BREAST
FILE REFERENCE: 5972, US, P2
CURRENT APPLICATION NUMBER: US/09/215,818A
EARLIER FILING DATE: 1998-12-18
EARLIER APPLICATION NUMBER: 08/912,276
EARLIER FILING DATE: 1997-08-17
EARLIER APPLICATION NUMBER: 08/697,105
EARLIER FILING DATE: 1996-08-19
EARLIER APPLICATION NUMBER: 08/912,149
EARLIER FILING DATE: 1997-08-15
EARLIER APPLICATION NUMBER: 08/697,106
EARLIER FILING DATE: 1996-08-19
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 93
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-215-818-5

Query Match 12.6%; Score 12; DB 4; Length 93;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKLWVLMAL 12
DB 1 MKLWVLMAL 12

RESULT 9
PCT-US96-08235-2
Sequence 2, Application PC/TUS9608235
GENERAL INFORMATION:
APPLICANT: WATSON, MARK A.
APPLICANT: FLEMING, TIMOTHY P.
TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: ROGERS, HOWELL & HAFERKAMP
STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
CITY: ST. LOUIS
STATE: MISSOURI
COUNTRY: USA
ZIP: 63105-1817
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/08235
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 964796
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314) 727-5188
TELEFAX: (314) 727-6092
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 93 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
PCT-US96-08235-2

Query Match 12.6%; Score 12; DB 5; Length 93;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLWVLMAL 12
DB 1 MKLWVLMAL 12

RESULT 10
US-08-403-852D-21
Sequence 21, Application US/08403852D
Patent No. 5891695
GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanche, Francis
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibault, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent

APPLICANT: De Crecy-Lagard, Valerie
TITLE OF INVENTION: Polypeptides Involved In The
Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852D
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 528 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-403-852D-21

Query Match
Best Local Similarity 7.4%; Score 7; DB 2; Length 528;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LAALLH 15
Db 501 LAALLH 507

RESULT 11
US-08-510-646B-22
Sequence 22, Application US/08510646B
Patent No. 6077699
GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanche, Francis
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
APPLICANT: De Crecy-Lagard, Valerie
TITLE OF INVENTION: Polypeptides Involved In The
Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.

COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,646B
FILING DATE: 03-AUG-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,852
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 528 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-510-646B-22

Query Match
Best Local Similarity 7.4%; Score 7; DB 3; Length 528;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LAALLH 15
Db 501 LAALLH 507

RESULT 12
US-09-231-818-21
Sequence 21, Application US/09231818
Patent No. 6171846
GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanche, Francis
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagorec, Monique
APPLICANT: Debussche, Laurent
APPLICANT: De Crecy-Lagard, Valerie
TITLE OF INVENTION: Polypeptides Involved In The
Biosynthesis Of Streptogramins, Nucleotide Sequences
TITLE OF INVENTION: Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/231,818
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403,852
FILING DATE: 10-MAY-1995
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 528 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-231-818-21

Query Match 7.48; Score 7; DB 4; Length 528;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 LAALLH 15
|||||||
Db 501 LAALLH 507

RESULT 13
US-07-699-468-2
Sequence 2, Application US/07699468
Patent No. 5350574
GENERAL INFORMATION:
APPLICANT: Erlanger, Bernard F.
APPLICANT: Cleveland, William L.
APPLICANT: Cacalano, Nicholas A.
TITLE OF INVENTION: NOVEL DERIVATIVES OF CYCLOSPORINE A,
TITLE OF INVENTION: ANTIBODIES DIRECTED THERETO AND USES
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/699,468
FILING DATE: 19910513
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 280009
FILING DATE: 05-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 33654-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550

TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: AMINO ACID
TOPOLOGY: circular
MOLECULE TYPE: peptide
HYPOTHETICAL: N
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Calpha-methyl,trans-2-butadiene"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2
OTHER INFORMATION: /note= "Cdelta-methyl"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /note= "Methyl"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4
OTHER INFORMATION: /note= "Methyl"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /note= "Methyl"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 8
OTHER INFORMATION: /note= "Calpha-R"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note= "Methyl"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 10
OTHER INFORMATION: /note= "Methyl"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 11
OTHER INFORMATION: /note= "Methyl"
US-07-699-468-2

Query Match 6.38; Score 6; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 LAALL 14
|||||||
Db 6 LAALL 11

RESULT 14
US-07-699-468-3
Sequence 3, Application US/07699468
Patent No. 5350574
GENERAL INFORMATION:
APPLICANT: Erlanger, Bernard F.
APPLICANT: Cleveland, William L.
APPLICANT: Cacalano, Nicholas A.
TITLE OF INVENTION: NOVEL DERIVATIVES OF CYCLOSPORINE A,
TITLE OF INVENTION: ANTIBODIES DIRECTED THERETO AND USES
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York

STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/699,468
FILING DATE: 19910513
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 280009
FILING DATE: 05-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 33654-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: AMINO ACID
TOPOLOGY: circular
MOLECULE TYPE: peptide
HYPOTHETICAL: N
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note=
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /note="MeGly"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4
OTHER INFORMATION: /note="MeLeu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /note="MeLeu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 8
OTHER INFORMATION: /note="MeLeu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note="MeLeu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 10
OTHER INFORMATION: /note="MeLeu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 11
OTHER INFORMATION: /note="MeLeu"
US-07-699-468-3

Query Match 6.3%, Score 6; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 LAALLL 14
| | | | |
Db 6 LAALLL 11

RESULT 15
US-07-699-468-4
Sequence 4, Application US/07699468
Patent No. 5350574
GENERAL INFORMATION:
APPLICANT: Eplanger, Bernard F
APPLICANT: Cleveland, William L
APPLICANT: Cacalano, Nicholas A
TITLE OF INVENTION: NOVEL DERIVATIVES OF CYCLOSPORINE A,
ANTIBODIES DIRECTED THERETO AND USES
THEREOF
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/699,468
FILING DATE: 19910513
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 280009
FILING DATE: 05-DEC-1988
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 33654-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: AMINO ACID
TOPOLOGY: circular
MOLECULE TYPE: peptide
HYPOTHETICAL: N
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note=
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /note="MeGly"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4
OTHER INFORMATION: /note="MeLeu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 6
OTHER INFORMATION: /note="MeLeu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 8
OTHER INFORMATION: /note="MeLeu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note="MeLeu"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 11
OTHER INFORMATION: /note="MeLeu"

```

; NAME/KEY: Modified-site
; LOCATION: 10
; OTHER INFORMATION: /note= "MeLeu"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 11
; OTHER INFORMATION: /note= "MeLeu"
US-07-699-468-4

```

```

Query Match      6.38; Score 6; DB 1; Length 11;
Best Local Similarity 100.08; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 9 LAALLL 14
   |||||
Db 6 LAALLL 11

```

Search completed: January 25, 2003, 13:35:50
 Job time : 19 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2003, 13:35:25 : Search time 36 seconds
(without alignments)
351.634 Million cell updates/sec

Title: US-09-806-302a-2

Perfect score: 95

Sequence: 1 MKLLMLMLAALLLHCYADS.....NFGLMHTVDSINCKNKS 95

Scoring table:
Gapop 60.0 , Gapect 60.0
908470 seqs, 13325062 residues

Searched: 908470 seqs, 13325062 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	95	100.0	95	AAW35804	Human endometrial
2	95	100.0	95	AAO2590	A human mammoqlobl
3	95	100.0	95	AAO3769	Human endometrial
4	95	100.0	95	AAV92226	Human endometrial
5	95	100.0	95	AAV92227	Human endometrial
6	95	100.0	95	AAV63394	Mamoglobin homolo
7	95	100.0	95	AAV31682	Human 5' EST relat
8	95	100.0	95	AAO20555	An endometrial spe
9	95	100.0	95	ABO96635	Protein of human L
10	95	100.0	108	AAV60038	Human endometrial

11	73	76.8	74	19	AAW61649	Non-ocular disease
12	12	12.6	13	22	AAE51129	Human mammaglobin
13	12	12.6	90	22	AAE07528	Human mammaglobin
14	12	12.6	90	22	AAE07537	Human mammaglobin
15	12	12.6	93	18	AAW10179	Mammary-specific s
16	12	12.6	93	19	AAW59777	Amino acid sequenc
17	12	12.6	93	19	AAW48432	Mammaglobin protel
18	12	12.6	93	20	AAV01718	Mammaglobin, a mam
19	12	12.6	93	21	AAV13786	Human mammaglobin
20	12	12.6	93	21	AAV84622	Amino acid sequenc
21	12	12.6	93	22	AAE07517	Human mammaglobin
22	12	12.6	93	22	AAE07529	Human mammaglobin
23	12	12.6	93	22	AAE07530	Human mammaglobin
24	12	12.6	93	22	AAE07531	Human mammaglobin
25	12	12.6	93	22	AAE07532	Human mammaglobin
26	12	12.6	93	22	AAE07533	Human mammaglobin
27	12	12.6	93	22	AAE07535	Human mammaglobin
28	12	12.6	93	22	AAE07536	Human mammaglobin
29	12	12.6	93	22	AAE51127	Human mammaglobin
30	12	12.6	410	22	AAU33558	Human breast cance
31	12	12.6	743	22	AAU33558	Human breast cance
32	12	12.6	1095	22	AAU33557	Human breast cance
33	11	11.6	20	22	AAE51115	Human mammaglobin
34	11	11.6	20	22	AAE51133	Human mammaglobin
35	11	11.6	33	19	AAW48433	Mammaglobin synthe
36	11	11.6	74	21	AAE84624	Amino acid sequenc
37	11	11.6	93	22	AAE07534	Human mammaglobin
38	11	11.6	220	22	AAO22141	Ra12-mamaglobin f
39	10	10.5	10	22	AAE51142	Human mammaglobin
40	10	10.5	20	22	AAE51132	Human mammaglobin
41	9	9.5	9	22	AAE51116	Human mammaglobin
42	9	9.5	9	22	AAE51137	Human mammaglobin
43	9	9.5	9	22	AAE51138	Human mammaglobin
44	9	9.5	9	22	AAE51139	Human mammaglobin
45	8	8.4	20	19	AAW48436	Mammaglobin synthe

ALIGNMENTS

RESULT 1	
ID	AAW35804 standard; Protein: 95 AA.
XX	
AC	AAW35804:
XX	
DT	27-MAR-1998 (first entry)
XX	
DE	Human endometrial specific steroid-binding factor III.
XX	
KW	Endometrial specific steroid-binding factor III; ESF III; human;
KW	Clara cell secretory protein; endometrium;
KW	phospholipase A2 inhibitor; polychlorinated biphenyl; antiagregant;
KW	inflammation; asthma; rhinitis; cystic fibrosis; airway disease;
KW	neoplasia; atopy; therapy; diagnosis.
XX	
OS	Homo sapiens.
XX	
FH	
FT	Key
FT	Peptide
FT	Protein
FT	Location/Qualifiers
FT	1..21
FT	/label= sig_peptide
FT	22..95
FT	/label= Mat_protein
PN	MO9734997-A1.
XX	
PD	25-SEP-1997.
XX	
PE	21-MAR-1996;
XX	
PE	96WO-US03857.
XX	
PR	21-MAR-1996;
XX	
PR	96WO-US03857.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.

XX Gentz RL, Ni J, Yu G;
 XX WPI: 1997-480206/44.
 DR N-PSDB: AAT94832.
 XX
 PT Human endometrial specific steroid-binding factor I, II and III -
 PT or preventing disorders associated with expression of HMH
 PT airway disease, neoplasia, atopy etc.
 XX
 PS Claim 19; Page 65; 92pp; English.
 XX
 CC This sequence comprises human endometrial specific steroid binding
 CC factor III (ESF III), a protein that inhibits phospholipase A2
 CC activity, binds to polychlorinated biphenyl compounds, reduces
 CC foreign protein antigenicity, inhibits monocyte and neutrophil
 CC chemotaxis and phagocytosis, inhibits platelet aggregation
 CC regulates eicosanoid levels in the human uterus and controls the
 CC growth of endometrial cells. The amino acid sequence was deduced
 CC from a cDNA clone (see AAT94632) derived from a human endometrial
 CC tumour. ESF I (see AAW35802) and ESF II (see AAW35803) are also
 CC claimed. Human ESF III has about 36% identity with rat prostatic
 CC steroid-binding protein. Recombinant ESF I, II and III can be
 CC expressed in host cells for use in claimed methods (a) for treating
 CC a patient in need of ESF I, II or III (including expression of the
 CC polypeptide in vivo) and (b) for identifying compounds which bind
 CC to and inhibit activation of the ESF polypeptide. hESF I, II and
 CC III may be used to treat inflammation, asthma, rhinitis, cystic
 CC fibrosis, airway disease, neoplasia and atopy.
 CC
 SQ Sequence 95 AA;

Query Match 100.0%; Score 95; DB 18; Length 95;
 Best Local Similarity 100.0%; Pred. No. 1.5e-89;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLWVLMALALLHCHYADSCCKLEDMVEKTSIDSIPEYKELLQEFIDSDAAAEAMG 60
 DB 1 MRLWVLMALALLHCHYADSCCKLEDMVEKTSIDSIPEYKELLQEFIDSDAAAEAMG 60
 OY 61 KFKQCFLNQSHRTLNFGMLMHTVYDSIWCNMKS 95
 DB 61 KFKQCFLNQSHRTLNFGMLMHTVYDSIWCNMKS 95

RESULT 2
 AAY02590
 ID AAY02590 standard; Protein: 95 AA.
 XX
 AC AAY02590;
 XX

DT 26-JUL-1999 (first entry)
 XX
 DE A human mamaglobin homologue (HMH).
 XX

KW Human mamaglobin homologue; HMH; antagonist; neoplastic disorder;
 KW adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma;
 KW teratocarcinoma; endometriosis.
 XX

OS Homo sapiens.
 XX

PN W09919487-A1.
 XX

PD 22-APR-1999.
 XX

PF 14-OCT-1998; 98MO-US21729.
 XX

PR 16-OCT-1997; 97US-0951750.
 XX

PA (INCY-) INCYTE PHARM INC.
 XX

PI Hillman JL, Murry LE, Shah P;
 XX

DR WPI: 1999-302531/25.
 DR N-PSDB: AAX36138.
 XX
 PT New human mamaglobin homolog (HMH), useful for diagnosing, treating
 PT or preventing disorders associated with expression of HMH
 PT
 XX
 PS Claim 1; Fig 1A-B; 63pp; English.
 XX
 CC The present sequence represents a human mamaglobin homologue (HMH).
 CC Antagonists of the HMH polypeptide can be used to treat neoplastic
 CC disorders including adenocarcinoma, leukemia, lymphoma, melanoma,
 CC myeloma, sarcoma and teratocarcinoma. A vector expressing the
 CC complement of the polynucleotide encoding HMH may be administered
 CC to a subject to treat or prevent neoplastic disorders or endometriosis.
 CC Antibodies which bind HMH may also be used in the diagnosis of
 CC conditions or diseases characterized by expression of HMH, or in assays
 CC to monitor patients being treated with HMH, agonists, antagonists or
 CC inhibitors. Polynucleotides encoding HMH may also be used diagnostically
 CC to detect and quantitate gene expression in biopsied tissues. With
 CC respect to cancer a relatively high amount of transcript may indicate a
 CC predisposition for the development of disease. The nucleic acid sequences
 CC which encode HMH may also be used to generate hybridization probes useful
 CC for mapping the naturally occurring genomic sequence. HMH, and its
 CC fragments/variants can be used for screening libraries of compounds in
 CC drug screening techniques.
 CC
 SQ Sequence 95 AA;

Query Match 100.0%; Score 95; DB 20; Length 95;
 Best Local Similarity 100.0%; Pred. No. 1.5e-89;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRLWVLMALALLHCHYADSCCKLEDMVEKTSIDSIPEYKELLQEFIDSDAAAEAMG 60
 DB 1 MRLWVLMALALLHCHYADSCCKLEDMVEKTSIDSIPEYKELLQEFIDSDAAAEAMG 60
 OY 61 KFKQCFLNQSHRTLNFGMLMHTVYDSIWCNMKS 95
 DB 61 KFKQCFLNQSHRTLNFGMLMHTVYDSIWCNMKS 95

RESULT 3
 AAB03769
 ID AAB03769 standard; Protein: 95 AA.
 XX
 AC AAB03769;
 XX

DT 06-OCT-2000 (first entry)
 XX

DE Human endometrial specific steroid-binding factor III protein sequence.
 XX

KW Endometrial specific steroid-binding factor; human; hESF; inflammation;
 KW asthma; rhinitis; cystic fibrosis; air way disease; neoplasia; atopy;
 KW eicosanoid level regulator; chemotaxis inhibitor; endometrial cancer; ss.
 XX

OS Homo sapiens.
 XX

PN US6066724-A.
 XX

PD 23-MAY-2000.
 XX

PF 21-MAR-1997; 97US-0821451.
 XX

PR 21-MAR-1996; 96US-0014724.
 XX

PA (HUMA-) HUMAN GENOME SCI INC.
 XX

PI Yu G, Gentz R, Ni J;
 XX

DR WPI: 2000-375600/32.
 XX

DR N-PSDB: AAA59730.
 XX

PT Novel gene encoding human endometrial specific steroid-binding factor

PT	I, II and III which is useful for treating asthma, rhinitis, cystic
PR	fibrosis, airway disease and neoplasia
PS	Claim 1; Fig 3; 36pp: English.
XX	
CC	This invention relates to nucleic acid molecules encoding portions of the
CC	human endometrial specific steroid-binding factors I, II, and III. Also
CC	included in the invention are hsf I, II, and III polypeptide sequences.
CC	The nucleotide sequence exhibit antisthmatic, antiinflammatory,
CC	antiallergic, and cyostatic properties. The polynucleotides are used in
CC	gene therapy to express hsf I, II and III polypeptides in vivo to treat
CC	and/or prevent inflammation, asthma, rhinitis, cystic fibrosis, air way
CC	disease, neoplasia and atopy. The polynucleotides are also used to
CC	inhibit phospholipase A2 activity, bind polychlorinated biphenyls, reduce
CC	foreign protein antigenicity, inhibit monocyte and neutrophil chemotaxis
CC	and phagocytosis, inhibit platelet aggregation, regulate eicosanoid
CC	levels in the human uterus and control the growth of endometrial cells.
CC	The polynucleotides are also useful for detecting complementary
CC	polynucleotides as a diagnostic reagent. The hsf I, II and III
CC	polynucleotides are used to detect complementary polynucleotides such as
CC	a diagnostic reagent. Detection of a mutated form of hsf I, II and III
CC	associated with a dysfunction will provide a diagnostic tool that can
CC	define diagnosis of a disease or susceptibility to a disease which
CC	results from under-expression, over-expression or altered expression of
CC	hsf I, II and III e.g. a susceptibility to inherited asthma and
CC	endometrial cancer. They are also useful for chromosome identification.
CC	The present sequence represents a hsf III protein sequence identified in
CC	the invention.
XX	
SQ	Sequence 95 AA;
	Query Match 100.0%; Score 95; DB 21; Length 95;
	Best Local Similarity 100.0%; Pred. No. 1.5e-89;
	Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 MLLMLVLMALILHCYADSGCKLLEDWEKTIINSISIPYKELLOEFTIDSDAAAEAWG 60
DB	1 MLLMLVLMALILHCYADSGCKLLEDWEKTIINSISIPYKELLOEFTIDSDAAAEAWG 60
OY	61 KFKOCFLNOSHRTLNKFGMLMHVYDSDICCNKNSN 95
DB	61 KFKOCFLNOSHRTLNKFGMLMHVYDSDICCNKNSN 95
RESULT 4	
ID	AAV92226 standard; Protein: 95 AA.
AC	AAV92226
XX	
AC	AAV92226;
XX	
DT	10-AUG-2000 (first entry)
XX	
DE	Human endometrial specific steroid binding factor III.
XX	
KW	Endometrial specific steroid binding factor; ESBPIII; diagnosis;
KW	gynaecological cancer; uterine; breast; endometrial; ovarian;
XX	antibody; gene therapy; cyostatic.
OS	
XX	Homo sapiens.
PM	WO200020044-A1.
XX	
PD	13-APR-2000.
XX	
PF	30-SEP-1999; 99WO-US22753.
XX	
PR	02-OCT-1998; 98US-0102743.
XX	
PA	(DIAD-) DIADEXUS LLC.
XX	
PI	Macina RA;
XX	
DR	WPI: 2000-303649/26.

XX	N-PSDB; AAA09104.
PT	Diagnosing, staging and monitoring gynecological cancer comprising
P7	using an elevated level of ESBP11 in a patient as an indicator of
P7	cancer.
XX	
PS	Claim 6; Page 28-29; 32pp; English.
XX	
CC	"The levels of human endometrial specific steroid binding factor
CC	(ESBP11) can be measured and compared to control levels and used to
CC	diagnose the presence of a gynaecological (uterine, breast, endometrial,
CC	or ovarian) cancer in a patient. ESBP11 levels can also be used to
CC	diagnose metastasis, to stage or monitor gynecological cancer.
CC	Antibodies specific for ESBP11 can be used to treat gynecological
CC	cancers.
XX	
SO	Sequence 95 AA:
Qy	
Db	
Qy	1 MKLLMVLMLALILHCYADSGCKLLEDWVEKTIINDISIPYEKELDGFIDSDAAAEAG 60
Db	1 MKLLMVLMLALLHLHYADSGCKLLEDWVEKTIINDISIPYEKELDGFIDSDAAAEAG 60
Qy	61 KFKOCFLNDSHRTLKNFGLMHTVYDSITWCNKSN 95
Db	61 KFKOCFLNDSHRTLKNFGLMHTVYDSITWCNKSN 95
RESULT 5	
AAY92237	
ID	AAV92237 standard; protein; 95 AA.
AC	AAV92237;
XX	
DT	10-AUG-2000 (first entry)
XX	
DE	Mammoglobin homologue from clone Mamm-X.
XX	
KW	Clone Mamm-X; mammaglobin; breast cancer; cytostatic; anti-HIV;
KW	immunosuppressive; antiallergic; antimetabolic; antiinflammatory;
KW	antidiabetic; antiarteriosclerotic; vasotropic; neuroprotective;
KW	nootropic; dermatological; tranquilizer; vulnerary.
XX	
OS	Homo sapiens.
XX	
PN	WO2000020447-A2.
PD	13-Apr-2000.
XX	
PF	06-OCT-1999; 99WO-US23294.
PR	06-OCT-1998; 98US-0103195.
PR	05-OCT-1999; 99US-0103195.
XX	
PA	(CURA-) CURAGEN CORP.
XX	
P1	Shinkets RA;
XX	
DR	WP1; 2000-303741/26.
DR	N-PSDB; AAA09118.
PT	Nucleic acids encoding polypeptides with synollin-like, claudin-like or
P7	Cytokine-like activity, useful for treating diseases including cancer,
P7	Alzheimer's and atherosclerosis
XX	
PS	Claim 23; Fig 6; 11pp; English.
CC	
CC	Clone Mamm-X encodes a polypeptide that is 100 percent identical to human
CC	Mammaglobin B precursor, a potential marker of breast cancer nodal
CC	metastasis. The sequences are useful for treatment of diseases such as

CC cancer, immune disorders, autoimmune disease, acquired immune deficiency
 CC syndrome (AIDS), transplant rejection, allergy, infection by a
 CC pathologic agent or organism, inflammatory disorders, arthritis, a
 CC hematopoietic disorder, a skin disorder, atherosclerosis, restenosis, a
 CC neurological disease, Alzheimer's disease, trauma, spinal cord injury
 CC and skeletal disorders.
 CC
 CC

Sequence 95 AA;

Query Match 100.0%; Score 95; DB 21; Length 95;
 Best Local Similarity 100.0%; Pred. No. 1.3e-89;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLMLVLAALLHCHYADSGCKLLEDVWEKTIINSISPEYKELLOEFIDSDAAAEAMG 60
 DB 1 MKLMLVLAALLHCHYADSGCKLLEDVWEKTIINSISPEYKELLOEFIDSDAAAEAMG 60
 OY 61 KFKOCFLNOSHRTLNKFGMLMHTVYDSIWCNMKSN 95
 DB 61 KFKOCFLNOSHRTLNKFGMLMHTVYDSIWCNMKSN 95

RESULT 6

AAV65394
 ID AAV65394 standard; protein; 95 AA.

AAV65394;

01-FEB-2000 (first entry)

Human 5' EST related polypeptide SEQ ID NO:1555.

Human: 5' EST; expressed sequence tag; secreted protein; diagnosis;
 gene therapy; chromosome mapping; upstream regulatory sequence;
 forensic; location; development; protein synthesis; stability;
 regulation; identification.

Homo sapiens.

MO9953051-A2.

21-OCT-1999.

09-APR-1999; 99MO-1B00712.

09-APR-1998; 98US-0057719.

28-APR-1998; 98US-0069047.

(GEST) GENSET.

Dumas Milne Edwards J, Duclert A, Giordano J;

WPI: 2000-038446/03.

N-PSDB; AA243008.

Novel secreted protein 5' expressed sequence tag sequences used in
 diagnostic, forensic, gene therapy, and chromosome mapping procedures

Claim 3; Page 018; 837pp; English.

AA242265 to AA243075 represent novel 5' expressed sequence tag (EST)
 sequences, corresponding to human secreted proteins. AAV64651 to
 CC AA65438 represent the EST-related proteins corresponding to AA242265 to
 CC AA43052. The 5' ESTs can be used for producing secreted human gene
 CC products. They can be used to identify and isolate 5' untranslated
 CC regions (UTRs) and upstream regulatory regions which control the
 CC location, development stage, rate, and quantity of protein synthesis, as
 CC well as stability of mRNA. The ESTs are also useful as probes for
 CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
 CC also be used in forensic procedures to identify individuals, or in
 CC diagnostic procedures to identify individuals having genetic diseases
 CC resulting from abnormal gene expression. The products may also be used in
 CC gene therapy protocols. The nucleic acids encoding signal peptides can be

CC used for directing extracellular secretion of a polypeptide or the
 CC insertion of a polypeptide into a membrane, or importing a polypeptide
 CC into a cell. The proteins encoded by the EST sequences may be useful in
 CC treating a variety of human conditions. Secreted proteins have
 CC therapeutic value, and the identification of new secreted proteins is
 CC valuable. AA242249 to AA242264 and AA64644 to AAV64650 represent
 CC sequences used in the exemplification of the present invention.
 CC
 CC

Sequence 95 AA;

Query Match 100.0%; Score 95; DB 21; Length 95;
 Best Local Similarity 100.0%; Pred. No. 1.3e-89;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLMLVLAALLHCHYADSGCKLLEDVWEKTIINSISPEYKELLOEFIDSDAAAEAMG 60
 DB 1 MKLMLVLAALLHCHYADSGCKLLEDVWEKTIINSISPEYKELLOEFIDSDAAAEAMG 60
 OY 61 KFKOCFLNOSHRTLNKFGMLMHTVYDSIWCNMKSN 95
 DB 61 KFKOCFLNOSHRTLNKFGMLMHTVYDSIWCNMKSN 95

RESULT 7

AAB31682
 ID AAB31682 standard; protein; 95 AA.

AAB31682;

30-APR-2001 (first entry)

An endometrial specific steroid binding factor III.

Human: endometrial specific steroid binding factor; hESF; hESF1; hESF1L;
 hESF1L; inflammation; asthma; rhinitis; cystic fibrosis; airway disease;
 neoplasia; atopy; phospholipase A2; polychlorinated biphenyl; chemotaxis;
 phagocytosis; platelet aggregation; eicosanoid; endometrial cell.

Homo sapiens.

US6174992-B1.

16-JAN-2001.

08-MAR-1999; 99US-0263810.

21-MAR-1996; 96US-0014724.

21-MAR-1997; 97US-0821451.

(HUMA-) HUMAN GENOME SCI INC.

Ni J, Yu G, Gentz R;

WPI: 2001-158477/16.

N-PSDB; AAF25214.

New human endometrial specific steroid binding factors, useful for
 treating and preventing inflammation, asthma, rhinitis, cystic
 fibrosis, airway disease, neoplasia and atopy

Claim 1; Fig 3; 36pp; English.

The present sequence represents a human endometrial specific steroid
 binding factor (hESF). The specification describes hESF1, hESF1L, and
 CC hESF1L. hESF1, II and III polypeptides, and polynucleotides encoding
 CC them are useful for treating and preventing inflammation, asthma,
 CC rhinitis, cystic fibrosis, airway disease, neoplasia and atopy,
 CC inhibiting phospholipase A2 activity, binding polychlorinated
 CC biphenyls, reducing foreign protein antigenicity, inhibiting monocyte

CC and neutrophil chemotaxis and phagocytosis, inhibiting platelet aggregation, regulating eicosanoid levels in the human uterus, and for controlling the growth of endometrial cells. hESF polypeptides and CC nucleotides are also useful for research, biological, clinical or therapeutic purposes.

XX Sequence 95 AA:

Query Match 100.0%; Score 95; DB 22; Length 95;
Best Local Similarity 100.0%; Pred. No. 1.5e-89;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLWVLMALLLHCYADSGCKLEDMVEKTIINSIPYKELQEPIDSDAAAEAMG 60
DB 1 MKLWVLMALLLHCYADSGCKLEDMVEKTIINSIPYKELQEPIDSDAAAEAMG 60
OY 61 KFKQCFLNQSHRTLNKFGMLMHTVYDSIMCNMKN 95
DB 61 KFKQCFLNQSHRTLNKFGMLMHTVYDSIMCNMKN 95

RESULT 8

AA020555
ID AA020555 standard; Protein; 95 AA.

AC AA020555;

DT 27-JUN-2002 (first entry)

XX Protein of human lipophilin C.

XX Immunogenic epitope; hormonally regulated organ; malignant tumour;
XX Lipophilin; human.

OS Homo sapiens.

PN US2002034739-A1.

XX 21-MAR-2002.

XX 07-JUL-1998; 98US-0110716.

XX 07-JUL-1998; 98US-0110716.

XX (LEHR/) LEHRER R I.

XX (ZHAO/) ZHAO C.

XX (GLAS/) GLASGOW B J.

XX PI Lehrer RI, Zhao C, Glasgow BJ;

XX WPI: 2002-338922/37.

XX N-PSDB: AAK99492.

XX Peptides having the sequence of human lipophilin A, B and C are

XX associated with carcinomas of hormonally regulated organs and are

XX useful in the diagnosis and prognosis of various cancers -

XX Claim 7; Fig 6; 22pp; English.

XX The invention relates to a peptide comprising the amino acid sequences of human lipophilin A, B, or C or its allelic variant or fragment comprising at least one immunogenic epitope, which is purified and isolated, and may have the N-terminal acylated and/or C-terminal amidated or be a fusion protein. Molecules of the invention are used in the diagnosis and prognosis of malignant tumours associated with hormonally regulated organs such as uterus, ovary, prostate, testis, breast, kidney and thymus. This sequence represents the human lipophilin C protein of the invention.

XX Sequence 95 AA:

Query Match 100.0%; Score 95; DB 23; Length 95;
Best Local Similarity 100.0%; Pred. No. 1.5e-89;

Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLWVLMALLLHCYADSGCKLEDMVEKTIINSIPYKELQEPIDSDAAAEAMG 60

DB 1 MKLWVLMALLLHCYADSGCKLEDMVEKTIINSIPYKELQEPIDSDAAAEAMG 60

OY 61 KFKQCFLNQSHRTLNKFGMLMHTVYDSIMCNMKN 95

DB 61 KFKQCFLNQSHRTLNKFGMLMHTVYDSIMCNMKN 95

RESULT 9

AB09635
ID AB09635 standard; Protein; 95 AA.

AC AB09635;

DT 29-MAY-2002 (first entry)

XX Human endometrial specific steroid-binding factor (hESF) III.

XX Human: endometrial specific steroid-binding factor; ESF;

XX prostatic steroid-binding protein; hESF I; hESF II; hESF III; asthma.

XX Homo sapiens.

OS Homo sapiens.

PN US6338948-B1.

XX 15-JAN-2002.

XX 30-MAY-2000; 2000US-0583169.

XX 21-MAR-1996; 96US-014724P.

XX 21-MAR-1997; 97US-0821451.

XX 08-MAR-1999; 99US-0263810.

XX (HUMA-) HUMAN GENOME SCI INC.

XX NI J, Yu G, Gentz R;

XX WPI: 2002-215019/27.

XX N-PSDB: ABL41783.

XX New antibody specific for human endometrial specific steroid-binding factor (hESF) III, useful for detecting hESF III protein in biological sample and to isolate or identify clones expressing the protein -

XX Disclosure; Fig 3; 36pp; English.

XX The present sequence represents a endometrial specific steroid-binding factor (hESF) III. The full length protein has a molecular weight of 8.10 kDa. The protein has homology to rat prostatic steroid-binding factor (hESF) III. Antibodies which bind hESF proteins, such as hESF I, hESF II, and hESF III are useful for isolating or to identify clones expressing the polypeptides or to purify the polypeptides by affinity chromatography. Agonists and antagonists of hESF proteins are useful for treating and/or preventing susceptibility to asthma.

XX Sequence 95 AA:

Query Match 100.0%; Score 95; DB 23; Length 95;
Best Local Similarity 100.0%; Pred. No. 1.5e-89;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKLWVLMALLLHCYADSGCKLEDMVEKTIINSIPYKELQEPIDSDAAAEAMG 60
DB 1 MKLWVLMALLLHCYADSGCKLEDMVEKTIINSIPYKELQEPIDSDAAAEAMG 60

QY 61 KFKOCFLNQSHTLKNFGMLMHTVYDSIMCNMKS 95
 DB 61 KFKOCFLNQSHTLKNFGMLMHTVYDSIMCNMKS 95

RESULT 10

AAV60038 standard; Protein: 108 AA.

AAV60038;

31-JAN-2000 (first entry)

Human endometrium tumour EST encoded protein 98.

Endometrium: human; tumour; cancer; anticancer; cytostatic; EST:
 treatment; uterine; gene therapy; expressed sequence tag.

Homo sapiens.

DEL9817948-A1.

21-OCT-1999.

17-APR-1998; 98DE-1017948.

17-APR-1998; 98DE-1017948.

(META-) METAGEN GES GENOMFORSCHUNG MBH.

Rosenblum A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E:

WPI: 1999-591957/51.

DR N-PSDB; AA242013.

New nucleic acid sequences expressed in uterine cancer tissues, and
 derived polypeptides, for treatment of uterine and endometrial cancer
 and identification of therapeutic agents

Claim 23; Page 314; 444pp; German.

This invention describes novel human nucleic acid (cDNA) sequences (A),
 that are highly expressed in uterine tumour tissue and which have
 anticancer and cytostatic activity. (A) are used (i) for recombinant
 expression of polypeptides (B) and (ii) to isolate complete genes. (B)
 are used (i) to identify agents suitable for treatment of uterine or
 endometrial cancer; (ii) directly for treating these forms of cancer
 (including expression from gene therapy vectors) and (iii) for
 generation of specific antibodies. (A) are identified by assembling ESTs
 (expressed sequence tags) from a particular tissue type before comparison
 of expression patterns. This allows a significantly longer fragment of
 the gene to be revealed, so should reduce the number of failures
 associated with the fact that ESTs from different libraries may represent
 different parts of the same unknown gene, distorting the estimated
 frequency of occurrence in a particular tissue. AAV5941-Y60328 represent
 protein fragments encoded by the human endometrium tumour cDNA library
 derived EST fragments represented in AAV41981-242121.

Sequence 108 AA;

Query Match 100.0%; Score 95; DB 20; Length 108;

Best Local Similarity 100.0%; Pred. No. 1.7e-89;

Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLWLVLAALLHGYADSGCKLEDMVEKTSNDISIPYKELLQEFIDSDAAAEAMG 60
 DB 14 MKLWLVLAALLHGYADSGCKLEDMVEKTSNDISIPYKELLQEFIDSDAAAEAMG 73

QY 61 KFKOCFLNQSHTLKNFGMLMHTVYDSIMCNMKS 95
 DB 74 KFKOCFLNQSHTLKNFGMLMHTVYDSIMCNMKS 108

RESULT 11

AAW61649 standard; peptide: 74 AA.

AAW61649;

27-OCT-1998 (first entry)

Non-ocular disease marker 3.

Human; non-ocular disease; tear; cancer; breast; prostate.

Homo sapiens.

WO9835229-A1.

13-AUG-1998.

06-FEB-1998; 98WO-A000071.

07-FEB-1997; 97AU-0005009.

(MACO-) MACQUARIE RES LTD.

(UNIX) UNISEARCH LTD.

Bolls S, Gooley AA, Herbert B, Molloy M, Morris C;

Walsh B, Willcox M, Williams KL;

WPI: 1998-447373/38.

Screening for non-ocular disease - by analysing tears for marker
 proteins, particularly indicative of cancer and genetic disease.

also new proteins and nucleic acid encoding them

Claim 6; Page 9; 14pp; English.

The markers AAW61647-W61649 are used for screening for, or detecting,
 non-ocular disease by analysing tears. Biochemicals, specifically
 proteins, are isolated from tears, particularly by chromatography or
 electrophoresis, especially two-dimensional polyacrylamide gel
 electrophoresis (2D-PAGE), then detected, e.g. with labelled specific
 reagents, in (radio)immunoassay. The method is used to detect cancer,
 particularly of breast or prostate, or a genetic disease, in humans or
 animals.

Sequence 74 AA;

Query Match 76.8%; Score 73; DB 19; Length 74;

Best Local Similarity 100.0%; Pred. No. 4.2e-67;

Matches 73; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 DSGCKLLEDMVEKTSNDISIPYKELLQEFIDSDAAAEAMGKFKOCFLNQSHTLKNFG 78
 DB 1 DSGCKLLEDMVEKTSNDISIPYKELLQEFIDSDAAAEAMGKFKOCFLNQSHTLKNFG 60

QY 79 LMMHTVYDSIMCN 91
 DB 61 LMMHTVYDSIMCN 73

RESULT 12

AAAB51129 standard; Peptide: 13 AA.

AAAB51129;

20-MAR-2001 (first entry)

Human mamaglobin peptide SEQ ID NO:29.

Human; mamaglobin; breast cancer; detection; diagnosis; antibody;
 vaccine; cytostatic; antimamaglobin.


```

XX OS Homo sapiens.
XX XX WO200073338-A1.
XX PN 07-DEC-2000.
XX PD 26-MAY-2000; 2000WO-US14845.
XX PF 28-MAY-1999; 99US-0136528.
XX PR 01-JUN-1999; 99US-0137048.
XX XX
XX PA (CORI-) CORIXA CORP.
XX PA (HEND/) HENDRICKSON R. C.
XX PA (HOUG/) HOUGHTON R. L.
XX PA (REED/) REED S. G.
XX PI Fanger GR.
XX DR WPI; 2001-049928/06.
XX PT Polypeptide comprising at least seven consecutive amino acid residues
XX PT of human mamaglobin, useful in the treatment and detection of breast
XX PT cancer -
XX PS Example 4; Fig 10; 109pp; English.
XX CC The present invention describes human mamaglobin peptides (I)
XX CC comprising at least 7 consecutive residues. Also described are: (1) a
XX CC vaccine comprising (I) with an immunostimulant which is an adjuvant;
XX CC (2) an isolated antibody (Ab1) or its antigen-binding fragment, which
XX CC specifically binds to a mamaglobin epitope having the sequence of
XX CC Pro2-3; (3) an isolated antibody (Ab2) or its antigen-binding fragment
XX CC that specifically binds to glycosylated mamaglobin; (4) a method for
XX CC inhibiting the development of breast cancer in a patient, comprising
XX CC administering (1) or Ab1 or Ab2; and (5) a method (M1) for determining
XX CC the presence or absence of breast cancer in a patient, comprising
XX CC cytostatic activity. The polypeptides and antibodies are used in
XX CC vaccines for the prevention and treatment of breast cancer. They are
XX CC also used for diagnosis and monitoring of breast cancer. The present
XX CC sequence represents a human mamaglobin peptide sequence, which is
XX CC used in an example from the present invention.
XX SQ Sequence 13 AA.

Query Match 12.6%; Score 12; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.9e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLWVLMIAL 12
   |||||||
Db 1 MKLWVLMIAL 12

RESULT 13
AAE07528
ID AAE07528 standard; Protein; 90 AA.
XX AC
XX AC AAE07528;
XX XX
XX DT 06-NOV-2001 (first entry)
XX XX
XX DE Human mamaglobin 207A 2 61551.2 protein.
XX DE
XX KM Human; lipophilin B; cytostatic; vaccine; mamaglobin 207A 2 61551.2;
XX KM gene therapy; uteroglobin; cancer; breast; ovary; prostate.
XX OS Homo sapiens.
XX OS
XX PN WO200158947-A1.
XX PD 16-AUG-2001.
XX XX

```

```

PF 08-FEB-2001; 2001WO-US04439.
XX 11-FEB-2000; 2000US-0183495.
XX PR 28-JUN-2000; 2000US-0215735.
XX XX
XX PA (CORI-) CORIXA CORP.
XX PI Carter D, Vedvick TS, Vallieve-Douglass J, Houghton RL, Dillon DC;
XX DR WPI; 2001-497069/54.
XX DR N-PSDB; AAD13771.
XX XX
XX PT Novel isolated complex two lipophilin-like polypeptides linked by at
XX PT least one disulfide bond, used to treat or prevent breast, ovarian or
XX PT prostate cancer -
XX PS Example 5; Page 78-79; 91pp; English.
XX CC The invention relates to a complex comprising a lipophilin-like
XX CC polypeptide linked by at least one disulphide bond to a second
XX CC lipophilin-like polypeptide. Lipophilin-like protein are members of
XX CC uteroglobin superfamily. Lipophilin-like proteins are useful in the
XX CC preparation of vaccines. The complex containing lipophilin-like
XX CC proteins are useful for treating or preventing breast, ovarian or
XX CC prostate cancer. The complex is also used for determining the
XX CC presence or absence of cancer in a patient, or monitor the progress
XX CC of cancer in a patient. Lipophilin DNA is also useful in gene therapy.
XX CC The present sequence is human mamaglobin 207A 2 61551.2 which is a
XX CC lipophilin-like protein.
XX SQ Sequence 90 AA.

Query Match 12.6%; Score 12; DB 22; Length 90;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKLWVLMIAL 12
   |||||||
Db 1 MKLWVLMIAL 12

RESULT 14
AAE07537
ID AAE07537 standard; Protein; 90 AA.
XX AC
XX AC AAE07537;
XX XX
XX DT 06-NOV-2001 (first entry)
XX XX
XX DE Human mamaglobin variant #8.
XX DE
XX KM Human; lipophilin B; cytostatic; vaccine; mamaglobin; variant; mutant;
XX KM gene therapy; uteroglobin; cancer; breast; ovary; prostate; ss.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO200158947-A1.
XX PD 16-AUG-2001.
XX PF 08-FEB-2001; 2001WO-US04439.
XX PR 11-FEB-2000; 2000US-0183495.
XX PR 28-JUN-2000; 2000US-0215735.
XX PA (CORI-) CORIXA CORP.
XX PI Carter D, Vedvick TS, Vallieve-Douglass J, Houghton RL, Dillon DC;
XX DR WPI; 2001-497069/54.
XX DR N-PSDB; AAD13787.
XX XX

```

PT Novel isolated complex two lipophilin-like polypeptides linked by at
 PT least one disulfide bond, used to treat or prevent breast, ovarian or
 PT prostate cancer -
 XX
 XX Example 6; Page 88; 91pp; English.
 PS
 CC The invention relates to a complex comprising a lipophilin-like
 CC polypeptide linked by at least one disulfide bond to a second
 CC lipophilin-like polypeptide. Lipophilin-like proteins are members of
 CC the uteroglobin superfamily. Lipophilin-like proteins are useful in the
 CC preparation of vaccines. The complex containing lipophilin-like
 CC proteins are useful for treating or preventing breast, ovarian or
 CC prostate cancer. The complex is also used for determining the
 CC presence or absence of cancer in a patient, or monitor the progress
 CC of cancer in a patient. Lipophilin DNA is also useful in gene therapy.
 CC The present sequence is human mammaglobin variant which is obtained
 CC by deleting the amino acids 79-81. Mammaglobin is a lipophilin-like
 CC protein.
 CC
 XX
 SQ Sequence 90 AA:
 Query Match 12.6%; Score 12; DB 22; Length 90;
 Best Local Similarity 100.0%; Pred. No. 0.00015;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKLWVLMAL 12
 DB 1 MKLWVLMAL 12
 RESULT 15
 AAW10179
 ID AAW10179 standard; Protein: 93 AA.
 XX
 AC AAW10179;
 DT 12-AUG-1997 (first entry)
 XX
 DE Mammary-specific secretory protein, mammaglobin.
 XX
 KW mammaglobin; mammary-secretory protein; breast cancer; detection;
 KM neoplastic disease; diagnosis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT /label= signal_peptide
 FT 20..93
 FT Protein /label= mature_protein
 XX
 PN WO9638463-A1.
 XX
 PD 05-DEC-1996.
 XX
 PF 31-MAY-1996; 96WO-US08235.
 XX
 PR 31-MAY-1995; 95US-0455896.
 XX
 PA (UNIT) UNIV WASHINGTON.
 XX
 PI Fleming TP, Watson MA.
 XX
 DR WPI; 1997-034299/03.
 DR N-PSDB; AAT50925.
 XX
 PT Nucleic acid encoding mammary-specific secretory protein.
 PT mammaglobin - used to develop prods. for the early diagnosis and
 PT treatment of breast cancer neoplastic disease
 XX
 PS Claim 3; Fig 2; 54pp; English.
 CC The present sequence is that of a mammary-specific secretory protein

CC designated mammaglobin, which is overexpressed in 27% of stage I primary
 CC breast cancer tumours. The anonymous sequence tag previously designated
 CC DEST002 was used to demonstrate that mammaglobin is abundant in the
 CC breast cancer tumour cell line MDA-MB-415. To isolate the full-length
 CC mammaglobin cDNA (AAT50925), the mRNA was reverse transcribed from
 CC this cell line and cloned using the RACE PCR technique. The nucleic acid
 CC and protein can be used to develop prods. e.g. antibodies or probes, for
 CC the detection and treatment of breast neoplastic disease.
 CC
 XX
 SQ Sequence 93 AA:
 Query Match 12.6%; Score 12; DB 18; Length 93;
 Best Local Similarity 100.0%; Pred. No. 0.00016;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MKLWVLMAL 12
 DB 1 MKLWVLMAL 12

Search completed: January 25, 2003, 13:40:07
 Job time : 36 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 25, 2003, 13:38:25 : Search time 31 seconds
(without alignments)
631.435 Million cell updates/sec

Title: US-09-806-302a-2

Perfect score: 95
Sequence: 1 MKLMLVLAALHHCYADS.....NFCIAMHTVYDSIMCNKSN 95

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_21: *
1: sp._archaea: *
2: sp._bacteria: *
3: sp._fungi: *
4: sp._human: *
5: sp._invertebrate: *
6: sp._mammal: *
7: sp._mhc: *
8: sp._organelle: *
9: sp._phage: *
10: sp._plant: *
11: sp._rodent: *
12: sp._virus: *
13: sp._vertebrate: *
14: sp._unclassified: *
15: sp._virus: *
16: sp._bacteriophage: *
17: sp._archaeop: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	8.4	249	09Y4U0	09Y4U0 homo sapien
2	8	8.4	256	09RUM5	09RUM5 deinococcus
3	8	8.4	372	095601	095601 acomys 1gn1
4	8	8.4	352	08X357	08X357 escherichia
5	8	8.4	937	09YPI2	09YPI2 citrus tait
6	8	8.4	2105	091R74	091R74 apple stem
7	7	7.4	28	062677	062677 rattus norv
8	7	7.4	76	024155	024155 nicotiana t
9	7	7.4	92	09G207	09G207 neurospora
10	7	7.4	93	09GK61	09GK61 oryctolagus
11	7	7.4	107	09KNI4	09KNI4 vibrio chol
12	7	7.4	168	08W4T0	08W4T0 solanum phn
13	7	7.4	223	08TYO9	08TYO9 methanopyru
14	7	7.4	231	09AJX4	09AJX4 streptomyce
15	7	7.4	234	09JHY1	09JHY1 halobacteri
16	7	7.4	239	09RJV3	09RJV3 streptomyce

17	7	7.4	259	5	085MW6	085MW6 drosophila
18	7	7.4	259	2	0937S2	0937S2 bruceella me
19	7	7.4	291	10	09AMZ0	09AMZ0 oryza sativ
20	7	7.4	295	2	08VNT3	08VNT3 enterobacte
21	7	7.4	295	16	08YBR1	08YBR1 bruceella me
22	7	7.4	316	2	09KHW7	09KHW7 salmonella
23	7	7.4	354	5	095V40	095V40 plasmodium
24	7	7.4	358	16	0983D4	0983D4 rhizobium 1
25	7	7.4	384	4	09GZNO	09GZNO homo sapien
26	7	7.4	384	11	09EPB7	09EPB7 mus musculu
27	7	7.4	384	11	09ESP4	09ESP4 rattus norv
28	7	7.4	402	5	018279	018279 caenorhabdl
29	7	7.4	425	5	026458	026458 drosophila
30	7	7.4	425	5	024338	024338 drosophila
31	7	7.4	469	5	046183	046183 drosophila
32	7	7.4	497	2	093PA1	093PA1 microscilla
33	7	7.4	525	11	09PC15	09PC15 mus musculu
34	7	7.4	558	10	094E19	094E19 arabidopsis
35	7	7.4	572	1	059663	059663 pyrodicticum
36	7	7.4	576	16	09XA04	09XA04 streptomyce
37	7	7.4	590	6	002665	002665 pyrodicticum
38	7	7.4	598	10	09C8J0	09C8J0 arabidopsis
39	7	7.4	605	16	092X67	092X67 rhizobium m
40	7	7.4	627	10	065380	065380 arabidopsis
41	7	7.4	665	16	097SC2	097SC2 streptococ
42	7	7.4	693	10	09FZB6	09FZB6 arabidopsis
43	7	7.4	758	16	0910L8	0910L8 pseudomonas
44	7	7.4	767	10	094AJ9	094AJ9 arabidopsis
45	7	7.4	769	5	095YU0	095YU0 drosophila

ALIGNMENTS

RESULT 1

09Y4U0 PRELIMINARY: PRT: 249 AA.
AC 09Y4U0:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 29.1 kDa protein (Fragment).
GN DRFP3641052.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN.
RA Duesterhoeft A., Lauber J., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 MYB-LIKE DOMAIN.
DR EMBL: AL080063; CAB45694.1; -
DR InterPro: IPR001005; Myb_DNA_binding.
DR InterPro: IPR000433; ZnF_ZZ.
DR Pfam: PF00248; Myb_DNA-binding; 1.
DR Pfam: PF00369; ZnF_ZZ; 1.
DR SMART: SM00395; SANT; 1.
DR SMART: SM00291; ZnF_ZZ; 1.
DR PROSITE: PS50090; MYB_3; 1.
DR PROSITE: PS01357; ZF_ZZ_1; UNKNOWN_1.
KW Hypothetical protein; Nuclear protein.
FT NON_TER
SQ SEQUENCE 249 AA; 29104 MW; D8B874EA33FC32 CRC64:

Query Match 8.4%; Score 8; DB 4; Length 249;

Best Local Similarity 100.0%; Pred. No. 4.6;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0Y 40 PEYKELD 47
|||||||

```

DB 136 PEYKELQ 143

RESULT 2
Q9RUM5 PRELIMINARY: PRT; 256 AA.
AC Q9RUM5;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Outer membrane protein.
GN DRI359.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_Taxid=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI.
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.P., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Olt H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarewicz K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans RI."
RL Science 286:1571-1577(1999).
DR EMBL; AE001982; AAF10931.1; -.
DR TIGR; DR1359; -.
DR InterPro: IPR004872; Lipoprotein_9.
DR Pfam; PF03180; Lipoprotein_9; 1.
DR Complete proteome.
SQ SEQUENCE 256 AA; 27737 MW; CC72A304E1B33D0D CRC64;

Query Match 8.4%; Score 8; DB 16; Length 256;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LMLAALL 14
DB 5 LMLAALL 12

RESULT 3
Q956Q1 PRELIMINARY: PRT; 372 AA.
AC Q956Q1;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Cytochrome b (Fragment).
GN CYTB OR CYTOCHROME B.
OS Acomys ignitus (fairy spiny mouse).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eulheta; Rodentia; Sciurognathi; Muridae; Murinae; Acomys.
OX NCBI_Taxid=60745;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BUSHWACKERS;
RX MEDLINE=9834580; PubMed=9668005;
RA Barone P.O., Monnerot M., Gautun J.C.;
RT "Intra generic phylogeny of Acomys (Rodentia, Muridae) using
RT mitochondrial gene cytochrome b."
MO. Phylogenet. Evol. 9:560-566(1998).
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY

```

```

CC CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
DR EMBL; 296064; CAB09438.1; -.
DR InterPro: IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c1.
DR Pfam; PF00033; cytochrome_b_c1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; UNKNOWN_1.
DR PROSITE; PS00193; CYTOCHROME_B_OO; UNKNOWN_1.
DR Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
FT NON_TER 372 372
SQ SEQUENCE 372 AA; 41903 MW; 387EE9BAD5B94C6E CRC64;

Query Match 8.4%; Score 8; DB 8; Length 372;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LMLAALL 14
DB 234 LMLAALL 241

RESULT 4
Q8X5S7 PRELIMINARY: PRT; 552 AA.
AC Q8X5S7;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Formate-dependent nitrite reductase, possible assembly function
DE (Cytochrome c-type biogenesis protein NITE).
GN NRE OR Z5673 OR EGS5056.
OS Escherichia coli O157:H7.
OC Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_Taxid=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RT Nature 405:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=1128796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
DR EMBL; AE005640; AAC59272.1; -.
DR EMBL; AP002568; BAB38479.1; -.
DR InterPro: IPR002541; Cyt_c_biom.
DR InterPro: IPR003568; Cyt_c_biom.
DR InterPro: IPR003567; Cyt_c_biom.
DR Pfam; PF01578; Cyt_c_biom.
DR PRINTS; PR01410; CCB10GENESIS.
DR TIGRFAMs; TIGR00353; nite; 1.
DR Complete proteome.
SQ SEQUENCE 552 AA; 61368 MW; 849838B944EDB8AD CRC64;

Query Match 8.4%; Score 8; DB 16; Length 552;

```

Best Local Similarity 100.0%; Pred. NO. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 LMLALLL 14
|||||
DB 114 LMLALLL 121

RESULT 5

OYPI2 PRELIMINARY; PRT; 937 AA.
AC OYPI2;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ORF1 (Fragment).
OS Citrus tatter leaf virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Capilliovirus.
OX NCBI_TaxID=33759;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=94103780; PubMed=8277280;
RA Yoshikawa N., Imazumi M., Takahashi T., Inouye N.;
RT "Striking similarities between the nucleotide sequence and genome
RT organization of citrus tatter leaf and apple stem grooving
RT capillioviruses.";
RL J. Gen. Virol. 74:2743-2747(1993).
DR EMBL: D16368; BAA03869.1; -;
DR InterPro: IPR001788; RNA_dep_RNAPol2. |
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
FT NON_TER 1
SQ SEQUENCE 937 AA; 108705 MW; 286DBA9F829FAA3B CRC64;

Query Match 8.4%; Score 8; DB 12; Length 937;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 KELLDEFT 50
|||||
DB 55 KELLDEFT 62

RESULT 6

OYIR74 PRELIMINARY; PRT; 2105 AA.
AC OYIR74;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE 241K polypotein.
OS Apple stem grooving virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Capilliovirus.
OX NCBI_TaxID=28347;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=L1-23;
RC Teruchi H., Magome H., Yoshikawa N., Takahashi T.;
RT "Nucleotide sequence of the genome of apple stem grooving capilliovirus
RT 1 isolate and construction of an infectious cDNA clone of the genome
RT containing a c aniflower mosaic virus 35S RNA.";
RL Ann. Phytopathol. Soc. Jpn. 63:432-436(1997).
DR EMBL: AB004063; BAA98054.1; -;
DR MEROPS: C35.001; -;
DR InterPro: IPR002173; PFKB.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
DR Pfam: PF01443; Viral_helicase1.
DR PROSITE: PS00583; PFKB_KINASES_1; UNKNOWN_1.
KM Polypotein.
SQ SEQUENCE 2105 AA; 241787 MW; ABA9BA31633A1A4F CRC64;

Query Match 8.4%; Score 8; DB 12; Length 2105;

Best Local Similarity 100.0%; Pred. NO. 28;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 KELLDEFT 50
|||||
DB 1223 KELLDEFT 1230

RESULT 7

OY62677 PRELIMINARY; PRT; 28 AA.
AC OY62677; O09145; O09065;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Alpha-1 collagen type IV (Fragment).
GN COL4A1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eumetaria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97107443; PubMed=8950183;
RA Grande J.P., Meider D.C., Kluge D.L., Wieden E.D.;
RT "Structure of the rat collagen IV promoter.";
RL Biochim. Biophys. Acta 1309:85-88(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Grande J.P., Haugen J.D.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U85606; AAB47426.1; -;
KW Collagen.
FT NON_TER 28
SQ SEQUENCE 28 AA; 3113 MW; 96DFCACC741E07809 CRC64;

Query Match 7.4%; Score 7; DB 11; Length 28;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 LMLALLL 15
|||||
DB 13 LMLALLL 19

RESULT 8

OY24155 PRELIMINARY; PRT; 76 AA.
AC OY24155;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE TA20.
GN Nicotiana tabacum (Common tobacco).
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asterales; eunasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ANTHER.
RA Beals T.P., Goldberg R.B.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 53-76 FROM N.A.
RC TISSUE=ANTHER.
RA Koltunow A.M., Truettner J., Cox K.H., Wallroth M., Goldberg R.B.;
RL Plant Cell 2:1201-1224(1990).
RN [3]
RP SEQUENCE OF 53-76 FROM N.A.
RC TISSUE=ANTHER.
RA Beals T.P., Goldberg R.B.;

RL Submitted (NOV-1996) to the EMBL/GenBank/DBSJ databases.
 DR EMBL: U73164; AAB18190.1; -;
 DR EMBL: U73165; AAB18191.1; -;
 SQ SEQUENCE 76 AA: 8269 MW: 2F0M47BF7A52F7CC CRC64:

Query Match
 Best Local Similarity 100.0%; Score 7; DB 10; Length 76;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LMVLMLA 10
 |||||
 DB 19 LMVLMLA 25

RESULT 9

Q9C207 PRELIMINARY; PRT; 92 AA.

AC Q9C207: 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Hypothetical 10.6 kDa protein.
 GN 104H10.170.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;

SEQUENCE FROM N.A.
 RA Schulte U., Aign V., Hohelsel J., Brandt P., Fortmann B., Holland R.,
 RA Nykatura G., Mewes H.W., Mannhaupt G.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RL German Neurospora genome project;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBSJ databases.
 DR EMBL: AL513410; CAC28583.2; -;
 KW Hypothetical protein.
 SQ SEQUENCE 92 AA: 10642 MW: 548FEB43A8D72E98 CRC64:

Query Match
 Best Local Similarity 100.0%; Score 7; DB 3; Length 92;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KILWVLM 8
 |||||
 DB 58 KILWVLM 64

RESULT 10

Q9GK61 PRELIMINARY; PRT; 93 AA.

AC Q9GK61: 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Lipophilin CS.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eumetaria; Lagomorpha; Leporidae; Oryctolagus;
 OX NCBI_TaxID=9986;

SEQUENCE FROM N.A.
 RP TISSUE-SUBMAXILLARY;
 RA Zhao C., Nguyen T.X., Lehrer R.I.;
 RT "Rabbit Lipophilins";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBSJ databases.
 DR EMBL: AF308620; AAC42808.1; -;
 DR InterPro: IPR003627; Mamgb/prostatn.
 DR InterPro: IPR000329; Uteroglobln.
 DR Pfam: PF01099; Uteroglobln; 1.
 DR PRINTS: PRO0486; UTEROGLOBIN.
 DR ProDom: PD029354; Mamgb/prostatn; 1.

DR SMART: SM00996; UTR: 1
 SQ SEQUENCE 93 AA: 10609 MW: E2B015AD319B2249 CRC64:

Query Match
 Best Local Similarity 100.0%; Score 7; DB 6; Length 93;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 VLMALAL 12
 |||||
 DB 6 VLMALAL 12

RESULT 11

Q9KN14 PRELIMINARY; PRT; 107 AA.

AC Q9KN14: 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Hypothetical protein VCA0152.
 GN VCA0152.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;

SEQUENCE FROM N.A.
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Ermolaeva M.D., Vamathevan J., Bass S., Olin H., Dragoi I., Sellers P.,
 RA McDonald L., Uitterlind T., Fleischmann R.D., Niernan M.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 cholerae";
 RL Nature 406:477-483(2000).
 DR EMBL: AE004356; AAF96065.1; -;
 DR TIGR: VCA0152; -;
 DR InterPro: IPR005133; Phag_MbHG_YuFB.
 DR Pfam: PF03334; Phag_MbHG_YuFB.
 DR TIGRFAMs: TIGR01300; CPA3_mnhG_Phag; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 107 AA: 11488 MW: 3B0F0CC77BF2D9B8 CRC64:

Query Match
 Best Local Similarity 100.0%; Score 7; DB 16; Length 107;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 MLALLL 14
 |||||
 DB 3 MLALLL 9

RESULT 12

Q8W4T0 PRELIMINARY; PRT; 168 AA.

AC Q8W4T0: 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Resistance gene analog (Fragment).
 OS Solanum phureja x Solanum stenocum.
 OC Eukaryota; Viridiplantae; streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=176295;

SEQUENCE FROM N.A.
 RC STRAIN=CV. BD410-33;
 RA Simko I., Jones R.W.;
 RT "Mapping of Resistance Gene Analogs in Potato.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY059428; AL30114.1; -
 DR InterPro: IPR002182; NB-ARC.
 DR Pfam: PF00931; NB-ARC; 1.
 FT NON_TER 1 1
 FT NON_TER 168 168
 SQ SEQUENCE 168 AA: 19000 MW: DB34AF796AE74EC4 CRC64;
 Query Match 7.4%; Score 7; DB 10; Length 168;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 43 KELL0EF 49
 Db 41 KELL0EF 47
 RESULT 13
 ID 08TY09 PRELIMINARY; PRT: 223 AA.
 AC 08TY09;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 GN Predicted dinucleotide-binding enzyme.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 NCBI_TaxID=2320;
 [1]
 RN RN
 RC SEQUENCE FROM N.A.
 RC MEDLINE=21927647; PubMed=11930014;
 RA Slegarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
 RA Shecherdinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Nalete D.A., Rogozin I.B., Talusov R.L., Wolf Y.I., Stetter K.O.,
 RA Mal'kh A.G., Koonin E.V., Kozayakhin S.A.;
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 RT and monophyly of archaeal methanogens.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL: AF010321; AAM01450.1; -
 KW Complete proteome.
 SQ SEQUENCE 223 AA: 23975 MW: 3C379224BAF6F92 CRC64;
 Query Match 7.4%; Score 7; DB 17; Length 223;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 52 SDAAEA 58
 Db 64 SDAAEA 70
 RESULT 14
 ID 09AJX4 PRELIMINARY; PRT: 231 AA.
 AC 09AJX4;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 GN Putative oxidoreductase.
 GN SC01803 OR SC133.02.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=1902;
 RN RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Harris D.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RC MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
 RA Kinsash H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 [4]
 RN RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.F., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Lake L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RT Nature 417:141-147(2002).
 CC -I- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 DR EMBL: AL513407; CAC28543.1; -
 DR HSSP: 064105; 10AA.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PR00080; SDRFAMILY.
 DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
 KW Oxidoreductase.
 SQ SEQUENCE 231 AA: 23789 MW: 6AEDB4030694086 CRC64;
 Query Match 7.4%; Score 7; DB 16; Length 231;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 52 SDAAEA 58
 Db 139 SDAAEA 145
 RESULT 15
 ID 09HHY1 PRELIMINARY; PRT: 234 AA.
 AC 09HHY1;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 GN TRK potassium uptake system protein.
 GN TRK2 OR VNG6175C.
 OS Halobacterium sp. (strain NRC-1).
 OC Plasmid pNRC200.
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 NCBI_TaxID=64091;
 RN RN
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20504483; PubMed=11016950;
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Stroga J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir T.A., Goo Y.A.,
 RA Leitbauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Madocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spidlich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;

RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AE005152; AAC20842.1;-.
DR InterPro; IPR000309; TrkA_Kuptake.
DR InterPro; IPR003148; TrkA_N.
DR Pfam; PF02080; TrkA-C; 1.
DR Pfam; PF02254; TrkA-N; 1.
DR PRINTS; PRO0335; KUPTAKEPTRKA.
KW Plasmid; Complete proteome.
SQ SEQUENCE 234 AA; 25291 MW; 92D50B3E557ADFD CRC64;

Query Match 7.48; Score 7; DB 17; Length 234;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LNVLMIA 10
|||||
DB 80 LNVLMIA 86

Search completed: January 25, 2003, 13:41:17
Job time : 34 secs

XX (HUMA-) HUMAN GENOME SCI INC.
 XX Gentz RL, NI J, Yu G;
 XX WPI: 1997-480206/44.
 DR P-PSDB: AAW35804.
 XX Human endometrial specific steroid-binding factor I, II and III -
 PT used to treat inflammation, asthma, rhinitis, cystic fibrosis,
 XX airway disease, neoplasia, atopy etc.
 PS Claim 11: Page 64; 92pp; English.
 XX This polynucleotide, deposited as AWC 97403, encodes human
 CC endometrial steroid binding factor III (ESF III, see
 CC AAW35804), a protein that inhibits phospholipase A2 activity, binds
 CC to polychlorinated biphenyl compounds, reduces foreign protein
 CC antigenicity, inhibits monocyte and neutrophil chemotaxis and
 CC phagocytosis, inhibits platelet aggregation, regulates eicosanoid
 CC levels in the human uterus and controls the growth of endometrial
 CC cells. The polynucleotide was discovered in a cDNA library derived
 CC from a human endometrial tumour. ESF I (see AAT94830) and ESF II
 CC (see AAT94831) clones are also claimed. Human ESF III has about
 CC 36% identity with rat prostatic steroid-binding protein. ESF I, II
 CC and III polynucleotides can be used in the production of
 CC recombinant polypeptides in host cells, and in claimed methods for
 CC (a) the treatment of a patient in need of ESF I, II or III and (b)
 CC for the diagnosis of a disease or a susceptibility to a disease
 CC related to underexpression of ESF I, II or III. hESF I, II and III
 CC fibrosis, airway disease, neoplasia and atopy.

SQ Sequence 476 BP; 131 A; 115 C; 105 G; 125 T; 0 other:

Query Match 100.0%; Score 476; DB 18; Length 476;
 Best Local Similarity 100.0%; Pred. No. 1,6e-136;

Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 ACAGAGCTGCCAGCAGCATGTAACACACAGACAGCCGCTCGCATGAAGCTGTGATG 60
 QY 61 GTCTCATATGCTGGCGGCTCTCTCTGCACTGCTATGCAAGATTGCTCAAACTCTG 120
 DB 61 GTCTCATATGCTGGCGGCTCTCTCTGCACTGCTATGCAAGATTGCTCAAACTCTG 120
 QY 121 GAGGACATGTTGAAAGACATCAATTCGACATATCTATACCTGTAATCAAGAGCTT 180
 DB 121 GAGGACATGTTGAAAGACATCAATTCGACATATCTATACCTGTAATCAAGAGCTT 180
 QY 181 CTTCAAGAGTTTATAGACAGTGATCCGCTCAAGGCTATGGGAAATTCAGACAGTCT 240
 DB 181 CTTCAAGAGTTTATAGACAGTGATCCGCTCAAGGCTATGGGAAATTCAGACAGTCT 240
 QY 241 TTCTCTAACACAGTATGAGACTGTGAAAACTTTGGACATGATGATATAGATGTCAC 300
 DB 241 TTCTCTAACACAGTATGAGACTGTGAAAACTTTGGACATGATGATATAGATGTCAC 300
 QY 301 GACAGCATTTGCTGTAATATGAAAGTAAATTAACCTTAACCAAGCGCTTTGGCTCAGAC 360
 DB 301 GACAGCATTTGCTGTAATATGAAAGTAAATTAACCTTAACCAAGCGCTTTGGCTCAGAC 360
 QY 361 GCTACAGACTATAGCCAGACATCTGTGATTCGTAGAAACACTTTTCTTGTGTT 420
 DB 361 GCTACAGACTATAGCCAGACATCTGTGATTCGTAGAAACACTTTTCTTGTGTT 420
 QY 421 GCTTTTATGTGGGAAGCTGTAAGACACTGTTGAACCTCAATTCATTCATTTCA 476
 DB 421 GCTTTTATGTGGGAAGCTGTAAGACACTGTTGAACCTCAATTCATTCATTTCA 476

RESULT 2

AAAS9730
 ID AAAS9730 standard; cDNA; 476 BP.
 XX
 AC AAAS9730;
 DT 06-OCT-2000 (first entry)
 XX
 DE Human endometrial specific steroid-binding factor III coding sequence.
 XX
 KW Endometrial specific steroid-binding factor; human; hESF; inflammation;
 KW asthma; rhinitis; cystic fibrosis; air way disease; neoplasia; atopy;
 KW eicosanoid level regulator; chemotaxis inhibitor; endometrial cancer; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 46..333
 FT /tag= a
 FT /product= "hESF III"
 XX /note= "Endometrial specific steroid-binding factor III"
 XX
 PN US6066724-A.
 PD 23-MAR-2000.
 XX
 PF 21-MAR-1997; 97US-0821451.
 XX
 PR 21-MAR-1996; 96US-0014724.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Yu G, Gentz R, NI J;
 DR WPI: 2000-375600/32.
 DR P-PSDB: AAB03769.
 PT Novel gene encoding human endometrial specific steroid-binding factor
 PT I, II and III which is useful for treating asthma, rhinitis, cystic
 PT fibrosis, airway disease and neoplasia
 XX
 PS Claim 1; Fig 3; 36pp; English.
 XX
 CC This invention relates to nucleic acid molecules encoding portions of the
 CC human endometrial specific steroid-binding factors I, II, and III. Also
 CC included in the invention are hESF I, II, and III polypeptide sequences.
 CC The nucleotide sequence exhibit antiallergic, antiinflammatory,
 CC antiallergic, and cytosaric properties. The polynucleotides are used in
 CC gene therapy to express hESF I, II and III polypeptides in vivo to treat
 CC and/or prevent inflammation, asthma, rhinitis, cystic fibrosis, air way
 CC disease, neoplasia and atopy. The polynucleotides are also used to
 CC inhibit phospholipase A2 activity, bind polychlorinated biphenyls, reduce
 CC foreign protein antigenicity, inhibit monocyte and neutrophil chemotaxis
 CC and phagocytosis, inhibit platelet aggregation, regulate eicosanoid
 CC levels in the human uterus and control the growth of endometrial cells.
 CC The polynucleotides are also useful for detecting complementary
 CC polynucleotides as a diagnostic reagent. The hESF I, II and III
 CC polynucleotides are used to detect complementary polynucleotides such as
 CC a diagnostic reagent. Detection of a mutated form of hESF I, II and III
 CC associated with a dysfunction will provide a diagnostic tool that can
 CC define diagnosis of a disease or susceptibility to a disease which
 CC results from under-expression, over-expression or altered expression of
 CC hESF I, II and III e.g. a susceptibility to inherited asthma and
 CC endometrial cancer. They are also useful for chromosome identification.
 CC The present sequence represents a hESF III coding sequence identified in
 CC the invention.
 XX
 SQ Sequence 476 BP; 131 A; 115 C; 105 G; 125 T; 0 other:
 Query Match 100.0%; Score 476; DB 21; Length 476;
 Best Local Similarity 100.0%; Pred. No. 1,6e-136;
 Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACAGAGCTGCCAGCAGCATGTAACACACAGACAGCCGCTCGCATGAAGCTGTGATG 60

```

Db      1 ACGAGCTGCCAGCAGCATGSAACACAGACACACCCGCTCCGCTGCAAGCTGCTGATG 60
Oy      61 GTCTCTATGCTGGCGGCTCTCTCTGCACTGCTATGAGATTTGCTGCAAACTCTG 120
Db      61 GTCTCTATGCTGGCGGCTCTCTCTGCACTGCTATGAGATTTGCTGCAAACTCTG 120
Oy      121 GAGGACATGGTGTAAAAAGACATCATTCGACATATCTATACCTGAATACAAAGAGCTT 180
Db      121 GAGGACATGGTGTAAAAAGACATCATTCGACATATCTATACCTGAATACAAAGAGCTT 180
Oy      181 CTTCAAGATTTCAATAGCAAGTATGTCCTGCTGACAGGCTATGAGGAAATTCAGCAGTGT 240
Db      181 CTTCAAGATTTCAATAGCAAGTATGTCCTGCTGACAGGCTATGAGGAAATTCAGCAGTGT 240
Oy      241 TTCCCTCAACAGTCACATAGAACTGTGAATACTTGAGCTGATGATCATACAGTGTAC 300
Db      241 TTCCCTCAACAGTCACATAGAACTGTGAATACTTGAGCTGATGATCATACAGTGTAC 300
Oy      301 GACAGCATTTGGTGTATATGAAAGATTAATTAACCTTACCCAGGCGTTTGGCTCAGAGG 360
Db      301 GACAGCATTTGGTGTATATGAAAGATTAATTAACCTTACCCAGGCGTTTGGCTCAGAGG 360
Oy      361 GCTACAGACTATGAGCAGAACTATCTGTGATGCTGAGAAACCACTTTCTTGTGTT 420
Db      361 GCTACAGACTATGAGCAGAACTATCTGTGATGCTGAGAAACCACTTTCTTGTGTT 420
Oy      421 GCTTTTATGTGGGAAGTGTGTAGACAACTGTGAAACCTCAATTGATTCATTTC 476
Db      421 GCTTTTATGTGGGAAGTGTGTAGACAACTGTGAAACCTCAATTGATTCATTTC 476

RESULT 3
AAA09104
ID      AAA09104 standard: DNA; 476 BP.
XX
AC      AAA09104:
XX
DT      10-AUG-2000 (first entry)
XX
DE      Human endometrial specific steroid binding factor coding sequence.
XX
KW      Endometrial specific steroid binding factor; ESRP11; diagnosis;
KW      gynaecological cancer; uterine; breast; endometrial; ovarian;
KW      antibody; gene therapy; cytostatic; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FH      CDS
FT      46..333
FT      /*tag- a
XX
XX      WO200020044-A1.
XX
XX      13-APR-2000.
XX
XX      30-SEP-1999; 99MO-US22753.
XX
XX      02-OCT-1998; 98US-0102743.
XX
XX      (DIAD-) DIADEXUS LLC.
XX
XX      MacIna RA:
XX
XX      WPI; 2000-303649/26.
XX      DR      P-PSDB; AAY92226.
XX
XX      Diagnosing, staging and monitoring gynecological cancer comprising
XX      PT      using an elevated level of ESRP11 in a patient as an indicator of
XX      cancer
XX
XX      Claim 6; Page 28; 32pp; English.
XX

```

```

CC      The levels of human endometrial specific steroid binding factor
CC      (ESRP11) can be measured and compared to control levels and used to
CC      diagnose the presence of a gynecological (uterine, breast, endometrial,
CC      or ovarian) cancer in a patient. ESRP11 levels can also be used to
CC      diagnose metastasis, to stage or monitor gynecological cancer.
CC      Antibodies specific for ESRP11 can be used to treat gynecological
CC      cancers.
XX
SQ      Sequence 476 BP; 131 A; 115 C; 105 G; 125 T; 0 other:

Query Match      100.0%; Score 476; DB 21; Length 476;
Best Local Similarity 100.0%; Pred. No. 1,6e-136;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0:

Oy      1 ACGAGCTGCCAGCAGCATGSAACACAGACACACCCGCTCCGCTGCAAGCTGCTGATG 60
Db      1 ACGAGCTGCCAGCAGCATGSAACACAGACACACCCGCTCCGCTGCAAGCTGCTGATG 60
Oy      61 GTCTCTATGCTGGCGGCTCTCTCTGCACTGCTATGAGATTTGCTGCAAACTCTG 120
Db      61 GTCTCTATGCTGGCGGCTCTCTCTGCACTGCTATGAGATTTGCTGCAAACTCTG 120
Oy      121 GAGGACATGGTGTAAAAAGACATCATTCGACATATCTATACCTGAATACAAAGAGCTT 180
Db      121 GAGGACATGGTGTAAAAAGACATCATTCGACATATCTATACCTGAATACAAAGAGCTT 180
Oy      181 CTTCAAGATTTCAATAGCAAGTATGTCCTGCTGACAGGCTATGAGGAAATTCAGCAGTGT 240
Db      181 CTTCAAGATTTCAATAGCAAGTATGTCCTGCTGACAGGCTATGAGGAAATTCAGCAGTGT 240
Oy      241 TTCCCTCAACAGTCACATAGAACTGTGAATACTTGAGCTGATGATCATACAGTGTAC 300
Db      241 TTCCCTCAACAGTCACATAGAACTGTGAATACTTGAGCTGATGATCATACAGTGTAC 300
Oy      301 GACAGCATTTGGTGTATATGAAAGATTAATTAACCTTACCCAGGCGTTTGGCTCAGAGG 360
Db      301 GACAGCATTTGGTGTATATGAAAGATTAATTAACCTTACCCAGGCGTTTGGCTCAGAGG 360
Oy      361 GCTACAGACTATGAGCAGAACTATCTGTGATGCTGAGAAACCACTTTCTTGTGTT 420
Db      361 GCTACAGACTATGAGCAGAACTATCTGTGATGCTGAGAAACCACTTTCTTGTGTT 420
Oy      421 GCTTTTATGTGGGAAGTGTGTAGACAACTGTGAAACCTCAATTGATTCATTTC 476
Db      421 GCTTTTATGTGGGAAGTGTGTAGACAACTGTGAAACCTCAATTGATTCATTTC 476

RESULT 4
AAF25214
ID      AAF25214 standard: cDNA; 476 BP.
XX
AC      AAF25214:
XX
DT      30-APR-2001 (first entry)
XX
DE      cDNA encoding of an endometrial specific steroid binding factor III.
XX
XX      Human; endometrial specific steroid binding factor; hESF; hESF1;
XX      hESF11; inflammation; asthma; rhinitis; cystic fibrosis; airway disease;
XX      neoplasia; atopy; phospholipase A2; polychlorinated biphenyl; chemotaxis;
XX      phagocytosis; platelet aggregation; eicosanoid; endometrial cell; ss.
XX
XX      Homo sapiens.
XX
XX      OS
XX
XX      Key
XX      FH      Location/Qualifiers
XX      FT      46..333
XX      FT      CDS
XX
XX      /*tag- a
XX      /*product- "endometrial specific steroid binding
XX      /*tag- b
XX
XX      sig_peptide
XX
XX      US6174992-B1.
XX

```

```

XX 16-JAN-2001.
PD ABL41783
XX 08-MAR-1999; 9905-0263810.
PF 21-MAR-1996; 9605-0014724.
PR 21-MAR-1997; 9705-0821451.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX NI J, Yu G, Gentz R;
PI WPI: 2001-158477/16.
DR P-PSDB; AAB31682.
XX
XX New human endometrial specific steroid binding factors, useful for
PT treating and preventing inflammation, asthma, rhinitis, cystic
PT fibrosis, airway disease, neoplasia and atopy
XX
XX Example 1; Fig 3; 36pp; English.
PS
XX The present sequence encodes a human endometrial specific steroid
CC binding factor (hesf). The specification describes hesf, hesfII, and
CC hesfIII, hesfI, II and III polypeptides, and polynucleotides encoding
CC them are useful for treating and preventing inflammation, asthma,
CC rhinitis, cystic fibrosis, airway disease, neoplasia and atopy,
CC inhibiting phospholipase A2 activity, binding polychlorinated
CC biphenyls, reducing foreign protein antigenicity, inhibiting monocyte
CC and neutrophil chemotaxis and phagocytosis, inhibiting platelet
CC aggregation, regulating eicosanoid levels in the human uterus, and for
CC controlling the growth of endometrial cells. hesf polypeptides and
CC nucleotides are also useful for research, biological, clinical or
CC therapeutic purposes.
XX
XX Sequence 476 BP; 131 A; 115 C; 105 G; 125 T; 0 other:
SO
Query Match 100.0%; Score 476; DB 22; Length 476;
Best Local Similarity 100.0%; Pred. No. 1,6e-136;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAGAGCTGCACGACGACTGAAACAGACAGACAGCGCCCTCGCATGAAAGCTGTATG 60
DB 1 ACAGAGCTGCACGACGACTGAAACAGACAGACAGCGCCCTCGCATGAAAGCTGTATG 60
QY 61 GTCTCATGCTGGCGCCCTCTCTCTGCACTGCTATGCAAGTTGCGTGCAGAACTCTG 120
DB 61 GTCTCATGCTGGCGCCCTCTCTCTGCACTGCTATGCAAGTTGCGTGCAGAACTCTG 120
QY 121 GAGGACATGCTTGAAGAAGACATCAATTCGCATATCTATACCTGAATCAAGAGCTT 180
DB 121 GAGGACATGCTTGAAGAAGACATCAATTCGCATATCTATACCTGAATCAAGAGCTT 180
QY 181 CTTCAGAGCTTCAACAGAGTATGATGCGCTGCAGAGGCTATGGGAAATTCAGACAGT 240
DB 181 CTTCAGAGCTTCAACAGAGTATGATGCGCTGCAGAGGCTATGGGAAATTCAGACAGT 240
QY 241 TTCTCAACGACATGACATGAACTGTGAAAACCTTGAGCATGATGATACAGTGTAC 300
DB 241 TTCTCAACGACATGACATGAACTGTGAAAACCTTGAGCATGATGATACAGTGTAC 300
QY 301 GACAGCATTTGCTGTAAATATGAAGATAATTAACCTTACCCAGAGCGTTGGCTCAGAG 360
DB 301 GACAGCATTTGCTGTAAATATGAAGATAATTAACCTTACCCAGAGCGTTGGCTCAGAG 360
QY 361 GCTACAGACTATGGCAGAACTCATCTGTTGTTGCTGTGAAGAACCACTTCTCTGTGTT 420
DB 361 GCTACAGACTATGGCAGAACTCATCTGTTGTTGCTGTGAAGAACCACTTCTCTGTGTT 420
QY 421 GCTTTTATGTGGAACTGTAAGACACTGTGAAACCTCAATTCATTCATTCATTCATTC 476
DB 421 GCTTTTATGTGGAACTGTAAGACACTGTGAAACCTCAATTCATTCATTCATTCATTC 476

```

```

RESULT 5
ABL41783
ID ABL41783 standard; cDNA; 476 BP.
XX
XX ABL41783;
AC
XX 29-MAY-2002 (first entry)
DT
XX
XX Human endometrial specific steroid-binding factor (hesf) III cDNA.
DE
XX
XX Human; endometrial specific steroid-binding factor; ESF;
KW prostatic steroid-binding protein; hesf I; hesf II; hesf III; asthma; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FT CDS 46..333
FT FT /tag= a
FT FT /product= "endometrial specific steroid-binding factor
FT sig_peptide 46..108
FT FT /tag= b
FT FT mat_peptide 109..330
FT FT /tag= c
XX
XX US6338948-B1.
XX
XX 15-JAN-2002.
XX
XX 30-MAY-2000; 2000US-0583169.
XX
XX 21-MAR-1996; 9605-014724P.
PR 21-MAR-1997; 9705-0821451.
PR 08-MAR-1999; 9905-0263810.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX NI J, Yu G, Gentz R;
PI WPI: 2001-215019/27.
DR P-PSDB; ABB09635.
XX
XX New antibody specific for human endometrial specific steroid-binding
PT factor (hesf) III, useful for detecting hesf III protein in biological
PT sample and to isolate or identify clones expressing the protein
XX
XX Disclosure; Fig 3; 36pp; English.
PS
XX The present sequence encodes a endometrial specific steroid-binding
CC factor (hesf) III. The full length protein has a molecular weight of
CC 8.10 kDa. The protein has homology to rat prostatic steroid-binding
CC protein C3. Antibodies which bind hesf proteins, such as hesf I, hesf II,
CC and hesf III are useful for isolating or to identify clones expressing
CC the polypeptides or to purify the polypeptides by affinity
CC chromatography. Agonists and antagonists of hesf proteins are useful
CC for treating and/or preventing susceptibility to asthma.
XX
XX Sequence 476 BP; 131 A; 115 C; 105 G; 125 T; 0 other:
SO
Query Match 100.0%; Score 476; DB 24; Length 476;
Best Local Similarity 100.0%; Pred. No. 1,6e-136;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAGAGCTGCACGACGACTGAAACAGACAGACAGCGCCCTCGCATGAAAGCTGTATG 60
DB 1 ACAGAGCTGCACGACGACTGAAACAGACAGACAGCGCCCTCGCATGAAAGCTGTATG 60
QY 61 GTCTCATGCTGGCGCCCTCTCTCTGCACTGCTATGCAAGTTGCGTGCAGAACTCTG 120
DB 61 GTCTCATGCTGGCGCCCTCTCTCTGCACTGCTATGCAAGTTGCGTGCAGAACTCTG 120
QY 121 GAGGACATGCTTGAAGAAGACATCAATTCGCATATCTATACCTGAATCAAGAGCTT 180
DB 121 GAGGACATGCTTGAAGAAGACATCAATTCGCATATCTATACCTGAATCAAGAGCTT 180

```

```

Db 121 GAGGACATGGTTGAMAAAGACATTCATTCGACATATCTATACCTGAAATACAAAGACCTT 180
Qy 181 CTTCAAGAGTTCTATAGACAGTATGCCGCTGCAGAGCCTATGGGAAATTCACAGCAGTGT 240
Db 181 CTTCAAGAGTTCTATAGACAGTATGCCGCTGCAGAGCCTATGGGAAATTCACAGCAGTGT 240
Qy 241 TTCTCAACAGTCACATAGACTGTGAAAACTTTGGACGATGATGATCATACAGTGTAC 300
Db 241 TTCTCAACAGTCACATAGACTGTGAAAACTTTGGACGATGATGATCATACAGTGTAC 300
Qy 301 GACAGCATTGGTGTATATGAAGAGTAACTTTACCTTTACCCAGAGCCTTTGGCTCAGAGG 360
Db 301 GACAGCATTGGTGTATATGAAGAGTAACTTTACCTTTACCCAGAGCCTTTGGCTCAGAGG 360
Qy 361 GCTACAGACTATGGCCAGAACATCATCTGTTGATTTGCTGTGAAGAACCACTTTCTTGTGCTT 420
Db 361 GCTACAGACTATGGCCAGAACATCATCTGTTGATTTGCTGTGAAGAACCACTTTCTTGTGCTT 420
Qy 421 GCTTTTATGTGGAACTGCTAGACAACTGTTGAAACCTCAATTCATTCATTTCA 476
Db 421 GCTTTTATGTGGAACTGCTAGACAACTGTTGAAACCTCAATTCATTCATTTCA 476

RESULT 6
AAK9492
ID AAK9492 standard; cDNA: 496 BP.
AC AAK9492;
XX 27-JUN-2002 (first entry)
DE cDNA of human Lipophilin C.
XX Immunogenic epitope; hormonally regulated organ; malignant tumour;
KW Lipophilin; human; ss.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT primer_bind /tag= a
FT /note= "Bound to nucleotides 1-25 of primer P12"
FT CDS 41..328
FT /tag= b
FT /product= "Protein of human Lipophilin C"
FT primer_bind /tag= c
FT /note= "Bound to nucleotides 1-24 of primer P14"
FT primer_bind /tag= d
FT /note= "Bound to nucleotides 20-1 of primer P13"
FT primer_bind /tag= e
FT /note= "Bound to nucleotides 28-1 of primer P13"
FT modified_base /tag= f
FT /mod_base= OTHER
FT /note= "This nucleotide is featured as an X, which is not
FT further defined in the specification"
FT polyA_signal /tag= g
FT misc_difference /tag= h
FT /note= "This adenosine is optionally absent. The number
FT of adenosines is unspecified in the specification"
XX
PV US2002034739-A1.
PD 21-MAR-2002.
XX
PF 07-JUL-1998; 98US-0110716.
XX 07-JUL-1998; 98US-0110716.

```

```

XX (LEHR/) LEHRER R I.
PA (ZHAO/) ZHAO C.
PA (GLAS/) GLASGOW B J.
XX
PI Lehrer RI, Zhao C, Glasgow BJ;
DR WPI: 2002-338922/37.
DR P-PSDB; AAO20555.
XX
PT Peptides having the sequence of human lipophilin A, B and C are
PT associated with carcinomas of hormonally regulated organs and are
PT useful in the diagnosis and prognosis of various cancers
XX
PS Claim 8; Fig 6; 22pp; English.
XX
CC The invention relates to a peptide comprising the amino acid sequences of
CC human lipophilin A, B, or C or its allelic variant or fragment comprising
CC at least one immunogenic epitope, which is purified and isolated, and may
CC have the N-terminal acylated and/or C-terminal amidated or be a fusion
CC protein. Molecules of the invention are used in the diagnosis and
CC prognosis of malignant tumours associated with hormonally regulated
CC organs such as uterus, ovary, prostate, testis, breast, kidney and
CC thymus. This polynucleotide sequence represents the lipophilin C cDNA of
CC the invention.
XX
SO Sequence 496 BP; 141 A; 119 C; 102 G; 133 T; 1 other;
XX
Query Match 91.0%; Score 433.2; DB 24; Length 496;
Best Local Similarity 97.6%; Pred. No. 2.7e-123;
Matches 450; Conservative 0; Mismatches 9; Indels 2; Gaps 1;
Qy 6 CTGCCAGCAGCAGCTGACACAGACAGACGCGCGCTCCCATGAAAGCTGCTGATGGCTT 65
Db 1 CTGCCAGCAGCAGCTGACACAGACAGACGCGCGCTCCCATGAAAGCTGCTGATGGCTT 60
Qy 66 CATGCTGGCGGCCCTCCCTCTGCACTGCTATGACATTCCTGGCTGCAAACTCCGTGAGGA 125
Db 61 CATGCTGGCGGCCCTCCCTCTGCACTGCTATGACATTCCTGGCTGCAAACTCCGTGAGGA 120
Qy 126 CATGCTGAAAAGACCATCAATTCGACATATCTTACTGAAATCAAAAGGCTTCTTCA 185
Db 121 CATGCTGAAAAGACCATCAATTCGACATATCTTACTGAAATCAAAAGGCTTCTTCA 180
Qy 186 AGACTTCATAGACAGTATGCGCTCCACAGCCTATGGGAAATTCAAAGCAGTGTTCCT 245
Db 181 AGACTTCATAGACAGTATGCGCTCCACAGCCTATGGGAAATTCAAAGCAGTGTTCCT 240
Qy 246 CAACCACTCACATAGAACTCTGAAAACCTTTGGACTGTATGATGATACATGATGACAG 305
Db 241 CAACCACTCACATAGAACTCTGAAAACCTTTGGACTGTATGATGATGATGATGATGACAG 300
Qy 306 CATTTGCTATATATGAAGAGTAACTTATACCAAGCGCTTGGCTCAGAGGCTTAC 365
Db 301 CATTTGCTATATATGAAGAGTAACTTATACCAAGCGCTTGGCTCAGAGGCTTAC 360
Qy 366 AGACTATGCGCCAGACATCTGTGTATGCTGTAAGAACCACTTT--CTTCTGTGTGTCT 423
Db 361 AGACTATGCGCCAGACATCTGTGTATGCTGTAAGAACCACTTTCTTCTGTGTGTCT 420
Qy 424 TTTTATGTGGAACTGCTAGACAACTGTTGAAACCTCAAT 464
Db 421 TTTTATGTGGAACTGCTAGACAACTGTTGAAACCTCAAT 461

RESULT 7
AAZ42013
ID AAZ42013 standard; cDNA: 517 BP.
AC AAZ42013;
XX
DT 31-JAN-2000 (first entry)
XX

```

DE Human endometrium tumour cDNA derived EST 33.
 XX Endometrium: human: tumour: cancer: anticancer: cytostatic: EST:
 KW treatment: uterine: gene therapy: expressed sequence tag: ss.
 XX Homo sapiens.
 OS
 XX DE19817948-A1.
 PN
 XX 21-OCT-1999.
 PD
 XX 17-APR-1998: 98DE-1017948.
 PF
 XX 17-APR-1998: 98DE-1017948.
 PR
 XX 17-APR-1998: 98DE-1017948.
 PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
 PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pillarsky C, Dahl E;
 XX WPI: 1999-591957/51.
 DR P-PSDB: AAY60037, AAY60038, AAY60039.
 XX
 PT New nucleic acid sequences expressed in uterine cancer tissues, and
 PT derived polypeptides, for treatment of uterine and endometrial cancer
 PT and identification of therapeutic agents -
 PS
 XX Claim 3: Page 191: 444pp: German.
 CC This invention describes novel human nucleic acid (cDNA) sequences (A),
 CC that are highly expressed in uterine tumour tissue and which have
 CC anticancer and cytostatic activity. (A) are used (i) for recombinant
 CC expression of polypeptides (B) and (ii) to isolate complete genes. (B)
 CC are used (i) to identify agents suitable for treatment of uterine or
 CC endometrial cancer: (ii) directly for treating these forms of cancer
 CC (including expression from gene therapy vectors) and (iii) for
 CC generation of specific antibodies. (A) are identified by assembling ESTs
 CC (expressed sequence tags) from a particular tissue type before comparison
 CC of expression patterns. This allows a significantly longer fragment of
 CC the gene to be revealed, so should reduce the number of failures
 CC associated with the fact that ESTs from different libraries may represent
 CC different parts of the same unknown gene, distorting the estimated
 CC frequency of occurrence in a particular tissue. AA241981-242121 represent
 CC EST fragments derived from a human endometrium tumour cDNA library which
 CC encode the protein sequences represented in AAY59941-Y60328.
 CC
 XX
 XX Sequence 517 BP: 156 A; 121 C; 106 G; 134 T; 0 other:
 SQ
 Query Match 90.7%; Score 431.6; DB 20; Length 517;
 Best Local Similarity 97.4%; Pred. No. 8.5e-123;
 Matches 450; Conservative 0; Mismatches 9; Indels 3; Gaps 1;

DB 301 CATTGGCTGAATATAGAGTAATTAACCTTACCAGCGCTTGCTCAGAGGCTAC 360
 QY 366 AGACTATGGCCAGACTATCTGTGATGCTAGAAACACACTTCTTGTG---TTGC 422
 DB 361 AGACTATGCCAGACTATCTGTGATGCTAGAAACACACTTCTTGTGCTGTC 420
 QY 423 TTTTATGTCGGAGCTGTAGACAACTGTTGAACCTCAATT 464
 DB 421 TTTTATGTCGAANAACCTGTAGACAACTGTTGAACCTCAAAAT 462
 RESULT 8
 ID AAA09118 standard; cDNA: 517 BP.
 AC AAA09118:
 XX
 DT 10-AUG-2000 (first entry)
 DE
 DE Clone Mamm-X cDNA, encodes mamoglobin homologue.
 DE
 DE Clone Mamm-X: mamoglobin; breast cancer: cytostatic: anti-HIV;
 KW immunosuppressive: antiallergic: antiinfective: antiinflammatory:
 KW antirheptic: antirheosclerotic: vasotropic: neuroprotective;
 KW nootropic: dermatological; tranquilizer; vulnerary; ss.
 XX
 OS Homo sapiens.
 CC
 CC Key Location/Qualifiers
 CC CDS 65..352
 CC /tag= a
 XX
 XX M0200020447-A2.
 PD 13-APR-2000.
 PF 06-OCT-1999: 99MO-US23294.
 XX
 XX 06-OCT-1998: 98US-0103195.
 PR 05-OCT-1999: 99US-0103195.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Shinkets RA:
 PI
 DR WPI: 2000-303741/26.
 DR P-PSDB: AAY92237.
 XX
 PT NUCLEIC ACIDS encoding polypeptides with syncollin-like, claudin-like or
 PT cytokine-like activity, useful for treating diseases including cancer,
 PT Alzheimer's and atherosclerosis
 PS
 XX Claim 28: Fig 6; 11pp: English.
 CC Clone Mamm-X encodes a polypeptide that is 100 percent identical to human
 CC Mamoglobin B precursor, a potential marker of breast cancer nodal
 CC metastasis. The sequences are useful for treatment of diseases such as
 CC cancer, immune disorders, autoimmune disease, acquired immune deficiency
 CC syndrome (AIDS), transplant rejection, allergy, infection by a
 CC pathologic agent or organism, inflammatory disorders, arthritis, a
 CC hematopoietic disorder, a skin disorder, atherosclerosis, restenosis, a
 CC neurological disease, Alzheimer's disease, trauma, spinal cord injury
 CC and skeletal disorders.
 XX
 XX Sequence 517 BP: 144 A; 129 C; 105 G; 139 T; 0 other:
 SQ
 Query Match 90.7%; Score 431.6; DB 21; Length 517;
 Best Local Similarity 97.4%; Pred. No. 8.5e-123;
 Matches 450; Conservative 0; Mismatches 9; Indels 3; Gaps 1;

```

OY 66 CATGCTGGCGGCGCTCTCTGCTGCTGCTATGCAATTTCTGGCTGCAAACTCCTGGAGCA 125
DB 85 CATGCTGGCGGCGCTCTCTGCTGCTGCTATGCAATTTCTGGCTGCAAACTCCTGGAGCA 144
OY 126 CATGTTGAAAAGAACCATCAATTCGACATATCTATACCTGAATACAAAGAGCTTCTTCA 185
DB 145 CATGTTGAAAAGAACCATCAATTCGACATATCTATACCTGAATACAAAGAGCTTCTTCA 204
OY 186 AGACTATAGACAGTATGCGGCTGCAAGGCTATGCGGAATTCAGAGCTGTTTCT 245
DB 205 AGACTATAGACAGTATGCGGCTGCAAGGCTATGCGGAATTCAGAGCTGTTTCT 264
OY 246 CAACCACTACATAGACATCTGTAAAAAATTGGACTGATGATGATACAGTATGACAGAC 305
DB 265 CAACCACTACATAGACATCTGTAAAAAATTGGACTGATGATGATACAGTATGACAGAC 324
OY 306 CATTTGGTGAATATGAAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 365
DB 325 CATTTGGTGAATATGAAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 384
OY 366 AGACTATGCGGCGCTCTCTGCTGCTGCTATGCAATTTCTGGCTGCAAACTCCTGGAGCA 422
DB 385 AGACTATGCGGCGCTCTCTGCTGCTGCTATGCAATTTCTGGCTGCAAACTCCTGGAGCA 444
OY 423 TTTTATGTTGGGAAGTGTAGACAACTGTTGAAACCTCAATT 464
DB 445 TTTTATGTTGGGAAGTGTAGACAACTGTTGAAACCTCAAAAT 486

RESULT 9
AAZ43008
ID AAZ43008 standard; cDNA: 473 BP.
XX
AC AAZ43008:
XX
DT 01-FEB-2000 (first entry)
XX
DE Human 5' EST isolated from a cDNA library SEQ ID NO.167.
XX
KW Human; 5' EST: expressed sequence tag; secreted protein; diagnosis:
KW gene therapy; chromosome mapping; upstream regulatory sequence;
KW forensic; location; development; protein synthesis; stability;
KW regulation; identification; ss.
XX
OS Homo sapiens.
XX
PN M09953051-A2.
XX
PD 21-OCT-1999.
XX
PE 09-APR-1999; 99MO-IB00712.
XX
PR 09-APR-1998; 98US-0057719.
XX
PR 28-APR-1998; 98US-0069047.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J:
XX
DR WPI: 2000-018446/03.
XX
DR P-PSDB: AAY65394.
XX
PT Novel secreted protein 5' expressed sequence tag sequences used in
XX diagnostic, forensic, gene therapy, and chromosome mapping procedures
XX
PS Claim 1: Page 569: 837pp: English.
XX
CC AA42265 to AA43075 represent novel 5' expressed sequence tag (EST)
CC sequences, corresponding to human secreted proteins. AAY64651 to
CC AAY65438 represent the EST-related proteins corresponding to AA42265 to
CC AA43052. The 5' ESTs can be used for producing secreted human gene
CC products. They can be used to identify and isolate 5' untranslated

```

```

CC regions (UTRs) and upstream regulatory regions which control the
CC location, development stage, rate, and quantity of protein synthesis, as
CC well as stability of mRNA. The ESTs are also useful as probes for
CC chromosome mapping, and to obtain full length cDNA clones. The ESTs can
CC also be used in forensic procedures to identify individuals, or in
CC diagnostic procedures to identify individuals having genetic diseases
CC resulting from abnormal gene expression. The products may also be used in
CC gene therapy protocols. The nucleic acids encoding signal peptides can be
CC used for directing extracellular secretion of a polypeptide or the
CC insertion of a polypeptide into a membrane, or importing a polypeptide
CC into a cell. The proteins encoded by the EST sequences may be useful in
CC treating a variety of human conditions. Secreted proteins have
CC therapeutic value, and the identification of new secreted proteins is
CC valuable. AA42249 to AA42264 and AAY6464 to AAY64650 represent
CC sequences used in the exemplification of the present invention.
XX
SO Sequence 473 BP; 130 A; 113 C; 104 G; 123 T; 3 other.

Query Match 89.1%; Score 424; DB 21; Length 473;
Best Local Similarity 96.9%; Pred. No. 1.8e-120;
Matches 440; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

OY 6 CTGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 65
DB 20 CTGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 79
OY 66 CATGCTGGCGGCGCTCTCTGCTGCTGCTATGCAATTTCTGGCTGCAAACTCCTGGAGCA 125
DB 80 CATGCTGGCGGCGCTCTCTGCTGCTGCTATGCAATTTCTGGCTGCAAACTCCTGGAGCA 139
OY 126 CATGTTGAAAAGAACCATCAATTCGACATATCTATACCTGAATACAAAGAGCTTCTTCA 185
DB 140 CATGTTGAAAAGAACCATCAATTCGACATATCTATACCTGAATACAAAGAGCTTCTTCA 199
OY 186 AGACTATAGACAGTATGCGGCTGCAAGGCTATGCGGAATTCAGAGCTGTTTCT 245
DB 200 AGACTATAGACAGTATGCGGCTGCAAGGCTATGCGGAATTCAGAGCTGTTTCT 259
OY 246 CAACCACTACATAGACATCTGTAAAAAATTGGACTGATGATGATACAGTATGACAGAC 305
DB 260 CAACCACTACATAGACATCTGTAAAAAATTGGACTGATGATGATACAGTATGACAGAC 319
OY 306 CATTTGGTGAATATGAAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 365
DB 320 CATTTGGTGAATATGAAAGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 379
OY 366 AGACTATGCGGCGCTCTCTGCTGCTGCTATGCAATTTCTGGCTGCAAACTCCTGGAGCA 422
DB 380 AGACTATGCGGCGCTCTCTGCTGCTGCTATGCAATTTCTGGCTGCAAACTCCTGGAGCA 439
OY 423 TTTTATGTTGGGAAGTGTAGACAACTGTTGAA 456
DB 440 TTTTATGTTGGGAAGTGTAGACAACTGTTGAA 473

RESULT 10
AAZ436138
ID AAZ436138 standard; DNA: 485 BP.
XX
AC AAZ436138:
XX
DT 26-JUL-1999 (first entry)
XX
DE DNA encoding a human mammoglobin homologue (HMH).
XX
KW Human mammoglobin homologue; HMH; antagonist; neoplastic disorder;
KW adenocarcinoma; leukemia; lymphoma; melanoma; myeloma; sarcoma;
KW teratocarcinoma; endometrial; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 43..330

```

```
FT      /*tag= a
XX      /product= "mammothlobin homologue"
XX      W09919487-A1.
XX      22-APR-1999.
XX      14-OCT-1998; 98WO-US21729.
XX      16-OCT-1997; 97US-0951750.
XX      (INCY-) INCYTE PHARM INC.
XX      Hillman JL, Murry LE, Shah P;
XX      WPI; 1999-302531/25.
XX      P-PSDB; AAY02590.
XX      New human mammothlobin homolog (HMH), useful for diagnosing, treating
XX      or preventing disorders associated with expression of HMH
XX      Claim 7; Fig 1A-B; 63pp; English.
XX      The present sequence encodes a human mammothlobin homologue (HMH).
XX      Antagonists of the HMH polypeptide can be used to treat neoplastic
XX      disorders including adenocarcinoma, leukemia, lymphoma, melanoma,
XX      myeloma, sarcoma and teratocarcinoma. A vector expressing the
XX      complement of the polynucleotide encoding HMH may be administered
XX      to a subject to treat or prevent neoplastic disorders or endometriosis.
XX      Antibodies which bind HMH may also be used in the diagnosis of
XX      conditions or diseases characterized by expression of HMH, or in assays
XX      to monitor patients treated with HMH, agonists, antagonists or
XX      inhibitors. Polynucleotides encoding HMH may also be used diagnostically
XX      to detect and quantitate gene expression in biopsied tissues. With
XX      respect to cancer a relatively high amount of transcript may indicate a
XX      predisposition for the development of disease. The nucleic acid sequences
XX      which encode HMH may also be used to generate hybridization probes useful
XX      for mapping the naturally occurring genomic sequence. HMH, and its
XX      fragments/variants can be used for screening libraries of compounds in
XX      drug screening techniques.
XX      Sequence 485 BP; 136 A; 116 C; 102 G; 129 T; 2 other:
XX
XX      Query Match      85.8%; Score 408.4; DB 20; Length 485;
XX      Best Local Similarity 95.2%; Pred. No. 1,2e-115;
XX      Matches 455; Conservative 0; Mismatches 16; Indels 7; Gaps 3;
XX
XX      6 CTGGCAGCAGCAGTGAACAGACAGAGAGCCGCTCCGATGAGAGTGTGATGTCCT 65
XX      3 CTGGCAGCAGCAGTGAACAGACAGAGAGCCGCTCCGATGAGAGTGTGATGTCCT 62
XX
XX      66 CAGTGGTGGGGGCTCTCTCTGCTGCTGATGAGATGCTGCTGCGCAAACTCCGAGGA 125
XX      63 CAGTGGTGGGGGCTCTCTCTGCTGCTGATGAGATGCTGCTGCGCAAACTCCGAGGA 122
XX
XX      126 CAGTGGTGGAAAGACATCAATTCGACATATCTTACCTGAATCAAGAAGCTTCTTCA 185
XX      123 CAGTGGTGGAAAGACATCAATTCGACATATCTTACCTGAATCAAGAAGCTTCTTCA 182
XX
XX      186 AGAGTTGATGAGAGTGAAGTGGCGGCTGAGAGCTATGGGGAATTCAGAGAGTTCCT 245
XX      183 AGAGTTGATGAGAGTGAAGTGGCGGCTGAGAGCTATGGGGAATTCAGAGAGTTCCT 242
XX
XX      246 CAGCAGTCAATGAGTCTGAAAACTTTGAGTGAATGATGATGATGATGATGATGATG 305
XX      243 CAGCAGTCAATGAGTCTGAAAACTTTGAGTGAATGATGATGATGATGATGATGATG 302
XX
XX      306 CATTGGTGAATATGAAAGTAAATTAATTACCAAGGGCTTGGCTGAGAGGGCTAC 365
XX      303 CATTGGTGAATATGAAAGTAAATTAATTACCAAGGGCTTGGCTGAGAGGGCTAC 362
XX
XX      366 AGACTATG---CCAGACTGATCTGTTGATGATGATGATGATGATGATGATGATGATG 419
XX      363 AGACTATG---CCAGACTGATCTGTTGATGATGATGATGATGATGATGATGATGATG 419
```

```
DB      363 AGACTATGCGCCGAGAACATCCATCTGTGATGATGATGATGATGATGATGATG 422
XX
XX      QY      420 TGCCTTTTATGTGGGAACCTGCTAGACACTGTTGAAACCTC-AATTCATTCATTTCA 476
XX      423 CTCCTTTTATGTGGGAACCTGCTAGACACTGTTGAAACCTC-AATTCATTCATTTCA 480
XX
XX      RESULT 11
XX      ABL67870
XX      ID      ABL67870 standard; DNA: 491 BP.
XX
XX      AC      ABL67870;
XX      DT      15-MAY-2002 (first entry)
XX
XX      DE      Ovary cancer related gene sequence SEQ ID NO:6207.
XX
XX      KW      Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
XX      KW      stomach; lung; prostate; pancreas; carcinoma; antitumour; carcinos;
XX      KW      cytosolic; gene therapy; antineoplastic; Wilms' tumour; adenocarcinoma;
XX      KW      gene; ds.
XX
XX      OS      Homo sapiens.
XX
XX      PN      W0200194629-A2.
XX
XX      PD      13-DEC-2001.
XX
XX      PF      30-MAY-2001; 2001WO-US10838.
XX
XX      PR      05-JUN-2000; 2000US-209473P.
XX      PR      05-JUN-2000; 2000US-209531P.
XX      PR      18-SEP-2000; 2000US-23133P.
XX      PR      18-SEP-2000; 2000US-233617P.
XX      PR      20-SEP-2000; 2000US-234009P.
XX      PR      20-SEP-2000; 2000US-234034P.
XX      PR      20-SEP-2000; 2000US-234052P.
XX      PR      22-SEP-2000; 2000US-234509P.
XX      PR      22-SEP-2000; 2000US-234567P.
XX      PR      25-SEP-2000; 2000US-234923P.
XX      PR      25-SEP-2000; 2000US-234924P.
XX      PR      25-SEP-2000; 2000US-235077P.
XX      PR      25-SEP-2000; 2000US-235082P.
XX      PR      25-SEP-2000; 2000US-235134P.
XX      PR      25-SEP-2000; 2000US-235280P.
XX      PR      26-SEP-2000; 2000US-235637P.
XX      PR      26-SEP-2000; 2000US-235638P.
XX      PR      27-SEP-2000; 2000US-235711P.
XX      PR      27-SEP-2000; 2000US-235720P.
XX      PR      27-SEP-2000; 2000US-235840P.
XX      PR      28-SEP-2000; 2000US-235863P.
XX      PR      28-SEP-2000; 2000US-236028P.
XX      PR      28-SEP-2000; 2000US-236032P.
XX      PR      28-SEP-2000; 2000US-236033P.
XX      PR      28-SEP-2000; 2000US-236034P.
XX      PR      28-SEP-2000; 2000US-236109P.
XX      PR      28-SEP-2000; 2000US-236111P.
XX      PR      29-SEP-2000; 2000US-236842P.
XX      PR      29-SEP-2000; 2000US-236881P.
XX      PR      02-OCT-2000; 2000US-237172P.
XX      PR      02-OCT-2000; 2000US-237173P.
XX      PR      02-OCT-2000; 2000US-237178P.
XX      PR      02-OCT-2000; 2000US-237294P.
XX      PR      02-OCT-2000; 2000US-237295P.
XX      PR      02-OCT-2000; 2000US-237316P.
XX      PR      03-OCT-2000; 2000US-237425P.
XX      PR      03-OCT-2000; 2000US-237598P.
XX      PR      03-OCT-2000; 2000US-237604P.
XX      PR      03-OCT-2000; 2000US-237606P.
XX      PR      03-OCT-2000; 2000US-237608P.
XX      PR      01-NOV-2000; 2000US-244867P.
XX      PR      01-NOV-2000; 2000US-245084P.
XX
```


XX (AVAL-) AVALON PHARM.
XX
PI Young PE, Augustus M, Carter KC, Edner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI: 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
PS Claim 1: SEQ ID 6207; 44pp: English.

ID	Accession	Standard	CDNA	BP
AA117037	AA117037	standard	CDNA	333 BP
XX	AC	AA117037		
XX	DT	07-DEC-2001	(first entry)	
XX	XX	Human breast cancer expressed polynucleotide 9494		
XX	XX	Human breast cancer; cell marker; cytostatic; ss.		
XX	XX	Homo sapiens		
XX	XX	MO200151628-A2		
XX	XX	19-JUL-2001		
XX	XX	10-JAN-2001	2001MO-US00798	
XX	XX	14-JAN-2000	2000US-0176077	
XX	XX	14-MAR-2000	2000US-0189167	
XX	XX	24-MAR-2000	2000US-0192099	
XX	XX	29-MAR-2000	2000US-0193480	
XX	XX	15-MAY-2000	2000US-0205230	
XX	XX	09-JUN-2000	2000US-0211315	
XX	XX	25-JUL-2000	2000US-0220534	
XX	XX	(MILL.)	MILLENNIUM PREDICTIVE MEDICINE INC.	
XX	XX	Lallie J, Xu Y, Wang Y, Steinmann K;		
XX	XX	WPI: 2001-451856/48		
XX	XX	New peptide useful as a marker for the diagnosis of breast cancer		
XX	XX	Claim 1: page 1697; 3695pp; English		
XX	XX	The invention relates to human breast cancer expressed polynucleotides		
XX	XX	(AI070544-AI26789) and methods of assessing whether a patient is the		
XX	XX	afflicted with breast cancer by examining the correlation between the		
XX	XX	expression of certain markers and the cancerous state of breast cells.		
XX	XX	The polynucleotides and encoded polypeptides are potential markers for		
XX	XX	detecting, diagnosing, monitoring, characterising treating and		
XX	XX	potentially preventing breast cancer. The polynucleotides and encoded		
XX	XX	polypeptides are also useful for isolating compounds with cytostatic		
XX	XX	activity.		
XX	XX	Sequence 333 BP; 88 A; 93 C; 79 G; 73 T; 0 other:		
XX	XX	Query Match	61.3%; Score 291.8; DB 22; Length 333;	
XX	XX	Best Local Similarity	99.3%; Pred. No. 8.8e-80;	
XX	XX	Matches 293; Conservative	0; Mismatches 2; Indels 0; Gaps	
XX	XX	0;		
XX	XX	6 CTGGCAGGCAAGCACTAACAACAGACAGACGCGCGCTCCCATGAAGCTGCTGCTCT	65	
XX	XX	32 CTGGCAGGCAAGCACTAACAACAGACAGACGCGCGCTCCCATGAAGCTGCTGCTCT	91	
XX	XX	66 CATGCTGGCGGCGCT	125	
XX	XX	92 CATGCTGGCGGCGCT	151	
XX	XX	126 CATGCTGGCGGCGCT	185	
XX	XX	152 CATGCTGGCGGCGCT	211	
XX	XX	186 AGATTCATAGACAGATGATGCGCGCTGACAGAGCTATGCGGAATTCGAAGAGCTTTTCT	245	
XX	XX	212 AGATTCATAGACAGATGATGCGCGCTGACAGAGCTATGCGGAATTCGAAGAGCTTTTCT	271	
XX	XX	246 CACCACTCATAGACATCTCTGAAAACTTTGGACTGATGATGATGATGATGATGATGATG	300	
XX	XX	272 CACCACTCATAGACATCTCTGAAAACTTTGGACTGATGATGATGATGATGATGATGATG	326	

```

RESULT 14
AAL08758
ID AAL08758 standard; cDNA: 370 BP.
XX
XX AAL08758;
AC
XX 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 1215.
DE
XX Human breast cancer: cell marker: cytosolic; ss.
XX
XX Homo sapiens.
OS
XX
XX M0200151628-A2.
PN
XX 19-JUL-2001.
PD
XX
XX 10-JAN-2001; 2001MO-US00798.
PE
XX
XX 14-JAN-2000; 2000US-0176077.
PR
XX 14-MAR-2000; 2000US-0189167.
PR
XX 24-MAR-2000; 2000US-0192099.
PR
XX 29-MAR-2000; 2000US-0193480.
PR
XX 15-MAY-2000; 2000US-0205230.
PR
XX 09-JUN-2000; 2000US-0211315.
PR
XX 25-JUL-2000; 2000US-0220534.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX
XX Little J, Xu Y, Wang Y, Steinmann K;
PI
XX
XX WPI; 2001-451856/48.
DR
XX
XX New peptide useful as a marker for the diagnosis of breast cancer -
PT
XX
XX Claim 1; Page 270; 3695pp; English.
PS
XX
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterizing treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
CC
CC
CC Sequence 370 BP; 93 A; 104 C; 94 G; 79 T; 0 other;
SQ
Query Match 61.3%; Score 291.8; DB 22; Length 370;
Best Local Similarity 99.3%; Pred.No. 9.3e-80;
Matches 293; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
0y 6 CTGGACGCGACGACTGACAGACAGACAGCGCGCCCTGCGCATGAAGCTGCTGATGGTCT 65
Db 69 CTGGACGCGACGACTGACAGACAGACAGCGCGCCCTGCGCATGAAGCTGCTGATGGTCT 128
0y 66 CATCTGTGGCGCCCTCTCTCTGCTGACACTGCTATGACAGATTTGGCTGCAAACTCTGGAGGA 125
Db 129 CATCTGTGGCGCCCTCTCTCTCTGCTGACACTGCTATGACAGATTTGGCTGCAAACTCTGGAGGA 188
0y 126 CATGCTTGAAGAAACCATCAATTTCCGACATATCTATACCTGAATACAAAGAGCTTTCTTA 185
Db 189 CATGCTTGAAGAAACCATCAATTTCCGACATATCTATACCTGAATACAAAGAGCTTTCTTA 248
0y 186 AGAGTTTATAGACAGTGCATGCCGCTGCACAGAGGCTATGGGAAATTCACAGAGCTTTCTT 245
Db 249 AGAGTTTATAGACAGTGCATGCCGCTGCACAGAGGCTATGGGGAATTCACAGAGCTTTCTT 308
0y 246 CAACGACATCACATAGAACTCTGAAAAAAGCTTTGGAAGTGTATGATGATACAGTATGAC 300

```


THIS PAGE BLANK (USPTO)


```
Db 361 GTTGATGCTAGAAACACATTTCTTCTGTGTCTTCTTTATGTGAACTGCTAGAC 420
QY 446 AACGTGTGAACCTCAATT 464
Db 421 AACGTGTGAACCTCAAAAT 439
RESULT 4
US-09-867-701-6508
: Sequence 6508, Application US/09867701
: Patent No. US2002013237A1
: GENERAL INFORMATION:
: APPLICANT: Aglate, Paul A.
: APPLICANT: Jones, Robert
: APPLICANT: Harlocker, Susan L.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
: FILE REFERENCE: 210121.497
: CURRENT APPLICATION NUMBER: US/09/867,701
: CURRENT FILING DATE: 2001-05-29
: NUMBER OF SEQ ID NOS: 10912
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 6508
: LENGTH: 368
: TYPE: DNA
: ORGANISM: Homo sapien
US-09-867-701-6508
Query Match 58.8%; Score 280; DB 10; Length 368;
Best Local Similarity 95.8%; Pred. No. 2.2e-76;
Matches 299; Conservative 0; Mismatches 10; Indels 3; Gaps 1;
QY 156 AACTATACCTGAATCAAGAGCTTCTTCAAGATTCATATGACAGATCCCGCTGCAGA 215
Db 2 AACTATACCTGAATCAAGAGCTTCTTCAAGATTCATATGACAGATCCCGCTGCAGA 61
QY 216 GGGTATGGGAAATTCAGAGAGTGTTCCTCAACAGTCACATAGAACTCTGAAAACTT 275
Db 62 GGGTATGGGAAATTCAGAGAGTGTTCCTCAACAGTCACATAGAACTCTGAAAACTT 121
QY 276 TGGAGTGAATGATGATACAGTGTAGAGAGATTTGGTAAATATGAAGATTAATTAAT 335
Db 122 TGGAGTGAATGATGATACAGTGTAGAGAGATTTGGTAAATATGAAGATTAATTAAT 181
QY 336 TTACCCAAAGCGCTTGGCTCAGAGGCTACAGACTATAGCCAGAACTCATCTGTTGATTG 395
Db 182 TTACCCAAAGCGCTTGGCTCAGAGGCTACAGACTATAGCCAGAACTCATCTGTTGATTG 241
QY 396 CTAGAAACACCTTCTCTCTGNG---TTGCTTTTATGTGGGAACCTGTAAGAACTGT 452
Db 242 CTAGAAACACCTTCTCTCTGNG---TTGCTTTTATGTGGGAACCTGTAAGAACTGT 301
QY 453 GAAACCTCAATT 464
Db 302 GAAACCTCAAAAT 313
RESULT 5
US-09-956-999-5
: Sequence 5, Application US/09956999
: Patent No. US20020064792A1
: GENERAL INFORMATION:
: APPLICANT: Lincoln, Stephen
: APPLICANT: Klingler, Tod M.
: APPLICANT: Au-Young, Janice
: APPLICANT: Tang, Y. Tom
: APPLICANT: Gould, Richard
: APPLICANT: Akerdrom, Ingrid E.
: APPLICANT: Sellhammer, Jeffrey J.
: APPLICANT: Hawkins, Phillip R.
: APPLICANT: Murry, Lynn E.
: APPLICANT: Deleane, Angelo M.
: APPLICANT: Levine, Wendy B.
```

```
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Goll, Surya K.
: APPLICANT: Altus, Christina M.
: APPLICANT: Bandman, Olga
: APPLICANT: Labrie, Samuel T.
: APPLICANT: Shah, Puri
: TITLE OF INVENTION: Database for Storage and Analysis of
: TITLE OF INVENTION: Full Length Sequences
: FILE REFERENCE: 6514-069CON
: CURRENT APPLICATION NUMBER: US/09/956,999
: CURRENT FILING DATE: 2001-09-19
: PRIOR APPLICATION NUMBER: 08/282,955
: PRIOR FILING DATE: 1995-07-29
: PRIOR APPLICATION NUMBER: 08/187,530
: PRIOR FILING DATE: 1994-01-27
: PRIOR APPLICATION NUMBER: 08/179,873
: PRIOR FILING DATE: 1994-01-11
: PRIOR APPLICATION NUMBER: 08/100,523
: PRIOR FILING DATE: 1993-08-03
: PRIOR APPLICATION NUMBER: 08/137,951
: PRIOR FILING DATE: 1993-10-14
: PRIOR APPLICATION NUMBER: 07/977,780
: PRIOR FILING DATE: 1992-11-19
: PRIOR APPLICATION NUMBER: 07/916,491
: PRIOR FILING DATE: 1992-07-17
: PRIOR APPLICATION NUMBER: 08/289,822
: PRIOR FILING DATE: 1994-08-12
: PRIOR APPLICATION NUMBER: 08/581,240
: PRIOR FILING DATE: 1995-12-29
: PRIOR APPLICATION NUMBER: 08/744,026
: PRIOR FILING DATE: 1996-11-05
: PRIOR APPLICATION NUMBER: 08/786,999
: PRIOR FILING DATE: 1997-01-23
: PRIOR APPLICATION NUMBER: 08/822,262
: PRIOR FILING DATE: 1997-03-20
: PRIOR APPLICATION NUMBER: 08/951,750
: PRIOR FILING DATE: 1997-10-16
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 5
: LENGTH: 495
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-956-999-5
Query Match 47.0%; Score 223.8; DB 10; Length 495;
Best Local Similarity 71.4%; Pred. No. 4.6e-59;
Matches 325; Conservative 0; Mismatches 122; Indels 8; Gaps 2;
QY 7 TGGCAGCGACGACGTAGACAGACAGACGCGCTGCGCATGAACTGCTGATGCTCTC 66
Db 7 TGGCAGCGCGACGTAGACAGACAGACAGCGCTCACCATGAACTGCTGATGCTCTC 66
QY 67 ATGCTGGGGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 126
Db 67 ATGCTGGGGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 126
QY 127 ATGCTGAAAAGACCATTAATTCGACATATCTATACCTGATGATATCAAGAGCTTCTTCAA 186
Db 127 GGTATTTCCAAAGCAATCAATTCACAAAGTGTCTAAGACTGAAATCAAAAGCTTCTCAA 186
QY 187 GAGTTCATAGACGATGAGCGGCTGCGAGAGCTATGGGAAATTCAGAGAGCTTCTCTC 246
Db 187 GAGTTCATAGACGATGAGCGGCTGCGAGAGCTATGGGAAATTCAGAGAGCTTCTCTC 246
QY 247 AACCACTCACAATGAACTCTGAAAACTTTGAGCTGATGATGCACTCACTGATGACAGC 306
Db 247 AACCAACGATGAAACCTGAGCAATGTGAGGTGTTATGCAATTAATATATGACAGC 306
QY 307 ATTGGTGAATATGAAAGATTAATTAATTTACCAAGAGGCTTGGCTGAGAGGCTTCA 366
Db 307 AGCTTTTGATTT-----ATTTAATCTTCTGCAAGAGCTTGGCTGACAGAACTGCA 360
```


Db 436 TTTTACTCAAACTACAAGACAATTGTTGAACCT 470

RESULT 8

```

US-09-934-054-11
Sequence 11, Application US/09934054
Patent No. US20020107385A1
GENERAL INFORMATION:
APPLICANT: Akcibiom, Ingrid E.
Hillman, Jennifer L.
Murry, Lynn E.
Goli, Surya K.
Hawkins, Phillip R.
TITLE OF INVENTION: BREAST TUMOR SPECIFIC PROTEINS
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: CA
COUNTRY: USA
ZIP: 94025-6936
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/934,054
FILING DATE: 21-Aug-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/747,547
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0077 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-934-054-11

```

Query Match	47.0%;	Score 223.8;	DB 10;	Length 503;
Best Local Similarity	71.4%;	Pred. No. 4.6e+59;		
Matches 325;	Conservative 0;	Mismatches 122;	Indels 8;	Gaps 2

Qy	7	TGCACGACAGCATTAACACAGACAGACGGCGCTGCCATGAAGTGGTATGGTCTC	66
Db	22	TGCACCGCGCATGACACCGACACCGACAGGCTCAACATGAAGTGGTATGGTCTC	81
Qy	67	ATGCTGGCGGCGCTCTCCTCTGACATGCTATGACAGATTCTGGCTGCACAACTCCTGAGGAC	125
Db	82	ATGCTGGCGGCGCTCTCTCCACAGATGCTGACGACGGCTCTGGCTGGCGCTTATTTGGAGAT	141
Qy	127	ATGCTTGAAGAAGCATTCAATTCGAGATATCTATATCCGTAATATACAAAGGCTTCTTCAA	186
Db	142	GTAATTTTCCACACATCATCAATCCACAGTGTCTTAAGACGTAATATACAAAGACTTCTTCAA	203
Qy	187	GAGTCATATGACAGATGATGCGGTGACAGAGGCTATGGGGAAATTCACGAGTGGTCTCTC	245
Db	202	GAGTCATATGACAGACAAATGCCATCAAAATGCCATGATGAATTTGAAGAAATGTTTCTT	263
Qy	247	AACCACTCACAATGAACTCTGAAAACTTTTGACTGATGATGCATCACTGTACACAGC	306

Db	262	AACCAACGAGTGAACCTCTGACCAATGTTGAAGTCTTTATGCAATTAATATATGACAGC	321
Qy	307	ATTGGGTGTAATATATAGAGTAATTAACCTTACCCAGAGCGTTTGGCTCAAGAGGCTTACA	366
Db	322	AGCTCTTTCGATT-----ATTTAACTTCTTGCAGAGCACTTTGGCTCACAGAAAGTCA	375
Qy	367	GACTATGGCCAACTCATCTGTTGATTCGTAGAAAC--CACCTTCTCTTGTGTGCTT	424
Db	376	GGGTATGAGGACAACCAACTACGAGATTCGTGCACAAACCAACCTTCTCTTCTATATGCT	435
Qy	425	TATATGCGAAGCTCTGACACACTGTTGAAACCT	459
Db	436	TTTACTACAACCTCAGACAAATTTGTTGAACCT	470

RESULT 9

```

US-09-825-301-73
Sequence 73, Application US/09825301
Patent No. US2002000978A1
GENERAL INFORMATION:
APPLICANT: Houghton, Raymond L.
APPLICANT: Dillon, David C.
APPLICANT: Molesh, David A.
APPLICANT: Xu, Jiangchun
APPLICANT: Zehentner, Barbara
APPLICANT: Persing, David H.
TITLE OF INVENTION: METHODS, COMPOSITIONS AND KITS FOR THE DETECTION
AND MONITORING OF BREAST CANCER
FILE REFERENCE: 210121.513
CURRENT APPLICATION NUMBER: US/09/825,301
CURRENT FILING DATE: 2001-04-02
NUMBER OF SEQ ID NOS: 77
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 73
LENGTH: 503
TYPE: DNA
ORGANISM: Homo sapiens
US-09-825-301-73

```

Query Match	46.7%	Score 222.2;	DB 10;	Length 503;
Best Local Similarity	71.2%;	Pred. NO. 1.4e-58;		
Matches 324;	Conservative -0;	Mismatches 123;	Indels 8;	Gaps 2;

OY	7	TGCGACGACGACATCAACACAGACAGCAGCGCGCTGCCATGAGCTGCTGATGGCTCC	66
Db	22	TGCGACCGCGACTGAACACGACACGACGACGCTCAACATGAAGTGTGATGGTCTC	81
OY	67	ATGCTGGCGGCCCTCTCTCTCTGACGTCTATGACAGATTTGCGTGCAGAACTCTGGAGAC	126
Db	82	ATGCTGGCGGCCCTCTCTCTCTGACGTCTATGACAGAGGCTTGCGCTGCCCTATTGGAGAT	141
OY	127	ATGGTTGAAGAACCATCATCAATTCGACATATCTATACCTGGAATATACAAAGAGCTCTTCAA	186
Db	142	GTAATTTCCAAAGCAGCATCAATTCACCAAGTGTCTTACAGCTGATATACAAAGAACTTCTCAA	201
OY	187	GAGTTCATAGACAGTGAATGCGCGCTGCAGAGGCTATGCGGAGAAATTCAGACAGTGTTCCTC	246
Db	202	GAGTTCATAGACAGCAATGCGCCACTACAAATGCAATGATGAATTAAGCAATGTTTTCTT	261
OY	247	AACCACTGCACATGAAAGCTCTGAAAAACCTTTGGATGATGATGATCAATGATAGACAGC	306
Db	262	AACCAAGCGATGAAACCTCTGACGAAATGTGAGTGTTCCTCAATTAATATACACAGC	321
OY	307	ATTTTGGTGAATATAGAGAGTAATTAACATTTACCCAGAGCGTTTGGCTCAGAGGCTACA	366
Db	322	AGCTCTTTGGATTTT-----ATTTTAATCTTTTGCAGAGCACTTTGGCTACAGAACTGCA	375
OY	367	GACATATGCGCAGACATCATCTGTTGATTTGGTAGAAG--CACTTCTCTCTGTGTGCTT	424
Db	376	GGGTATGTGTGAGAAACCAACATCTGCGATTTGCTGCAGAAACACACCTTCTCTTTCTATGTCT	433
OY	425	TTTATGTGGAACTGCTGACACACTGTTTGAACCT	459

Db 436 TTTTACTACAACCTACAAGACAATTGTTGAACCT 470

```

RESULT 10
US-09-778-320-43/c
: Sequence 43, Application US/09778320
: Patient No. US20010034052A1
GENERAL INFORMATION:
APPLICANT: Dillon, Davin C.
APPLICANT: Day, Craig H.
APPLICANT: Jiang, Yugu
APPLICANT: Houghton, Raymond L.
APPLICANT: Mitcham, Jennifer
APPLICANT: Wang, Tongtong
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
FILE REFERENCE: 210121.491C5
CURRENT APPLICATION NUMBER: US/09/778.320
CURRENT FILING DATE: 2001-02-06
NUMBER OF SEQ. ID NOS: 301
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 43
LENGTH: 467
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(467)
OTHER INFORMATION: n = A,T,C or G
US-09-778-320-43

```

Query Match	Similarity	41.0%	Score 195.2	DB 10	Length 467
Best local	Similarity	69.7%	Pred. No. 2,76-50		
Matches 294	Conservative	0	Mismatches 120	Indels 8	Gaps 2
QY	40	CTCCGCATGAGCTGCTGATGGTCTCTCATGCTGGCGGCCCTCTCTCTGACACTGCTATGCA	99		
Db	467	CTCCACCATGAGTGGCTGATGGTCTCTCATGCTGGCGGCCCTCTCTCCGACACTGCTACGCA	408		
QY	100	GATTCCTGGCGCAAACTCTCCGAGAGACATGGTGGAAAAGACATCATTCGACATATTC	159		
Db	407	GCGTCTGGCGTCCGCCCTTATTGAGCATGTGATTTCCAGACATTCATTCACAAAGTCT	348		
QY	160	ATACCTGATATACAAAGAGCTCTTCAAGATTTCAATAGACAGTGAATGCGCGTGGAGAGCT	219		
Db	347	AGACATCAATACAAAGAACCTTCTCAAGATTCATAGACGCAATGCGCACTACAAATATGCA	288		
QY	220	ATGGGGAATTCAGAGAGTGTCTCTCAACACAGTCACATAGAACTCTGGAAAACCTTATGCA	279		
Db	287	ATGATGATGATTCAGAGCAATGTTTTTTATACAAAGCATGAACTTGTGAGCAATGTTGAG	228		
QY	280	CTGATGATGCATATACAGTGTACGACAGCATTTGCTGTATATATGAGAGATTAATTAATCTTAC	339		
Db	227	GTTTATTAATTAATTAATATATATGACAGCATCTTTGTGATTT-----ATTTTAACCTTCT	174		
QY	340	CCAAAGCGTTTGGCTCAGAGGCTTACACACATATGCGCCGAACCTATCTGTGATGCTGAG	399		
Db	173	GCAAGACCTTTGGCTCAGAGAACTCAGAGGATATGGTGAAGAACCAAGCATGAGATGCTGCG	114		
QY	400	AAAC--CACATCTCTCTGTTGTTGGTTTTATGTGGAAAGCATGCTAGACAACTGTTGAAC	457		
Db	113	AAACGACACCTTCTCTTCTTATGTCTTTTACTACAAACTACAAAGCATTTGTGAAAC	54		
QY	458	CT 459			
Db	53	CT 52			

```

: GENERAL INFORMATION:
: APPLICANT: Dillon, Davin C.
: APPLICANT: Day, Craig H.
: APPLICANT: Jiang, Yugu
: APPLICANT: Houghton, Raymond L.
: APPLICANT: Mitcham, Jennifer
: APPLICANT: Wang, Tonglong
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Harlocker, Susan L.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER
: FILE REFERENCE: 210121.491C6
: CURRENT APPLICATION NUMBER: US/09/910,689
: CURRENT FILING DATE: 2001-07-20
: NUMBER OF SEQ. ID NOS: 307
: SOFTWARE: FASTSeq For Windows Version 4.0
: SEQ ID NO 43
: LENGTH: 467
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 242..263
: OTHER INFORMATION: n = A,T,C or G
: US-09-910-689-43

```

Query Match Similarity	41.0%	Score 195.2	DB 10	Length 467
Best Local Similarity	69.7%	Pred. No. 2.7e-50		
Matches 294	Conservative	0	Mismatches 120	Indels 8
			Gaps	2

Query	40	CTCCGCATGAGCTGCTGATGGTGCCTATGCTGAGCGGCCCTCTCTCTGCACTGCTATGCA	99
Db	467	CTCCACCATGAGTTGCTGATGCTGCTCATCTGCGCGGCCCTCTCCGACCTGCTACGCA	408
Query	100	GATTCGCGCGCAAACTCCTGAGGAGCATGTTGTTAAAAAGCATCATCTCCGACATATCT	159
Db	407	GCGCTTGCGTCCCTTATTGGAGAAATGTGATTTCCAGACATATCATTCACAAAGTCT	348
Query	160	ATACCTGCAATACAAAGAGCTTCTCAAGAGTTCAATAGACACAGTGTGCGCGTGCAGAGCT	219
Db	347	AAGACGTAATACAAAGAACTTCTCAAGATTCATAGACGCAATGCCACTACAAATGCGC	288
Query	220	ATGGGGAAATTCAGCAGTGTTCCTCAACGACACATATGAATCTCGAAAACTTTGCA	279
Db	287	ATGATGATCAATTGAGGAAGATGTTTTTTTATCCAAACGGATGAACCTTGTAGACATGTTGAG	228
Query	280	CTGATGATGCATACAGTGTACGACAGCATTTGGTGTAAATGAAAGACTAAATTAACCTTAC	339
Db	227	GTGTTTATCCAAATTAATATATATGACGACGACTTTGTGATTT-----ATTTAACTTCT	174
Query	340	CCAAAGCGTTTGGCTTCAGAGSGCTACAGACATAGTCCGCAAGATCATCTGTTATGCTGAG	399
Db	173	GCAAGACCTTTGGCTACAGAACTGCAGAGATATGGTAGAAACCAAGCTACGCAATTCCTGC	114
Query	400	AAAC--CACTTTCTCTGTTGTGTTGTTTATGTGGGAAAGCTGACAACTGTTGAAC	457
Db	113	AAACCAACACCTTCTCTTCTATGTCTTTTACTACAAACACTACAAAGCAATGTGTGAAC	54
Query	458	CT 459	
Db	53	CT 52	

Qy 458 CT 459
||
Db 53 CT 52

RESULT 11
US-09-910-689-43/c
; Sequence 43, Application US/09910689
Patent No. US0020081609A1

```

; APPLICANT: Wang, Tongtong
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Bennington, Angela Ann
; APPLICANT: Zehentner, Barbara
; APPLICANT: Fanger, Gary R.
; APPLICANT: Reiter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.491C7
; CURRENT APPLICATION NUMBER: US/10/010.742
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 307
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 43
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: 242, 263
; OTHER INFORMATION: n = A,T,C or G
US-10-010-742-43

```

```

Query Match          41.0%; Score 195.2; DB 12; Length 467;
Best Local Similarity 69.7%; Pred. No. 2.7e-50;
Matches 294; Conservative 0; Mismatches 120; Indels 8; Gaps 2;

```

```

OY 40 CTCGCCATGAAGCTCTGATGTCCTCATGCTGGCGCCCTCCTCGACATGCTATGCA 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 467 CTCACCATGAAGTTCGATGTCCTCATGCTGGCGCCCTCCTCGACATGCTATGCA 408
OY 100 GATTCGGCTGCAACCTCCTCGACATGCTGTAAGAACCTCAATCCGACATCT 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 407 GCGCTGGCTGGCCCTTTTGGAGATGTGATTTCCAGCAATCAATCCGACATGCT 348
OY 160 ATACTGAATACAAAGAGCTTCTTCAAGACTTCATAGACAGTATGCCGTCGAGAGCT 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 347 AAGACGATACAAAGAGCTTCTTCAAGACTTCATAGACAGTATGCCGTCGAGAGCT 288
OY 220 ATGGGGAATTCAGCAGCTTTCCTCAACAGTCACATGAGCTGAGAACTTGA 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 287 ATGATGATTAAGGATGTTTNTTACCAAGGATGAGCAATGTTGAG 228
OY 280 CTGATGATCATACAGTGTACGACAGATTTGGTATATGAGAAGTATTAATCTTAC 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 227 GTGTTATGCAATTAATATATGACAGCAGTCTTTGTATTT-----ATTTAATCTTCT 174
OY 340 CCAAGGCGTTTGGCTCAAGGGCTAAGAGATGCCGACATCTGTGTATGCTAG 399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 173 GCAGACCTTTGGCTCAAGAGCTGAGGGTATGTTGAGAACCGAGTGGATGCTGC 114
OY 400 AAAC--CACTTCTCTTGTCTGCTCTTTTATGTGGACATGCTAGACAACTGTGAAC 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 113 AAACGACACCTTCTCTTCTTATGTCTTTACTACAAACTACAGAACATTTGTAAC 54
OY 458 CT 459
    ||
DB 53 CT 52

```

```

RESULT 13
US-09-757-417-46
; Sequence 46, Application US/09757417
; Patent No. US2002082216A1
; GENERAL INFORMATION:
; APPLICANT: Fanger, Gary R.
; APPLICANT: Foy, Theresa M.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Reed, Steven G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.479C1

```

```

; CURRENT APPLICATION NUMBER: US/09/757,417
; CURRENT FILING DATE: 2001-01-08
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-757-417-46

```

```

Query Match          32.2%; Score 153.4; DB 10; Length 399;
Best Local Similarity 72.4%; Pred. No. 1.7e-37;
Matches 199; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

```

```

OY 46 ATGAGCTGCTGATGCTCTCATGCTGGCGCCCTCCTCGACATGCTATGCAATCT 105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 118 ATGAGCTGCTGATGCTCTCATGCTGGCGCCCTCCTCGACATGCTATGCAATCT 177
OY 106 GCGTCGCAACCTCTGAGACATGTTGAAAAGACATATCCGACATATCTATACCT 165
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 178 GCGTCGCCCTTATGAGAAATGTATTCACAGCAATCAATCCAACTGTCTAAGACT 237
OY 166 GATACAAAGAGCTTCTCAAGAGTATAGACAGTATGCGCTGACAGGCTATGGG 225
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 238 GATACAAAGAGCTTCTCAAGAGTATAGACAGCAATGCAATCCAAATGCAATAGT 297
OY 226 AATTCGACAGCTGTTCTCCTCAACGATCAGATAGAACTGAAAAAATTTGAGACTATG 285
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 298 GATTCGAAAGAGTGTTCCTTACCAAAAGGATCAACTCTGAGCAATGTAGTGTTT 357
OY 286 ATGACATAGCTAGACAGATTTGGTAAATAT 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 358 ATGCAATTAATATATGACAGCAATCTTTGTGATTT 392

```

```

RESULT 14
US-10-007-805-492
; Sequence 492, Application US/10007805
; Patent No. US20020150581A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jianshun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vestick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margalita
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 492
; LENGTH: 1233
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-007-805-492

```

```

Query Match          32.2%; Score 153.4; DB 12; Length 1233;
Best Local Similarity 72.4%; Pred. No. 3.1e-37;
Matches 199; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

```

```

OY 46 ATGAGCTGCTGATGCTCTCATGCTGGCGCCCTCCTCGACATGCTATGCAATCT 105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 ATGAGCTGCTGATGCTCTCATGCTGGCGCCCTCCTCGACATGCTATGCAATCT 60
OY 106 GCGTCGCAACCTCTGAGACATGTTGAAAAGACATATCCGACATATCTATACCT 165

```

```
Db      61 GGGTGGCCCTTATTGGACAATGTGATTTCCAGACAAATCAATCCACAAGTGTCTAAGACT 120
Qy      166 GAATACAAAGACCTTCTTCACAGAGTTCATAGACAGTATGCCGCTGCAGAGCTATGGGG 225
Db      121 GAATACAAAGACCTTCTTCACAGAGTTCATAGACAGCAATGCACTACAAATGCCATAGAT 180
Qy      226 AAATTCAAAGCACTGTTTCCTCCACAGCACTGACATGAACTGAAAACTTTGGACGATG 285
Db      181 GAATTGAAAGAACTGTTTCTTACCAAAAGCATGAAACTCTGAGCAATGTTGAGTGT 240
Qy      286 ATGCATACAGTGTACGACAGCATTTGGTAAATAT 320
Db      241 ATGCATTAATATATGACAGCAGTCTTGTGATTT 275
```

RESULT 15

```
US-10-007-805-491
: Sequence 491, Application US/10007805
: Patent No. US20020150581A1
: GENERAL INFORMATION:
: APPLICANT: Jiang, Yugu
: APPLICANT: Dillon, Davin C.
: APPLICANT: Mitcham, Jennifer L.
: APPLICANT: Xu, Jiangchun
: APPLICANT: Harlocker, Susan L.
: APPLICANT: Hepier, William F.
: APPLICANT: Henderson, Robert A.
: APPLICANT: Fanger, Gary R.
: APPLICANT: Vedyick, Thomas S.
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Durham, Margareta
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: FILE REFERENCE: 210121.470C10
: CURRENT APPLICATION NUMBER: US/10/007,805
: CURRENT FILING DATE: 2001-12-07
: NUMBER OF SEQ ID NOS: 593
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 491
: LENGTH: 2232
: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-007-805-491
```

```
Query Match      32.2% Score 153.4; DB 12; Length 2232;
Best Local Similarity 72.4%; Pred. No. 4.2e-37;
Matches 199; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
```

```
Qy      46 ATGAAGCTGCTGATGCTCTCATGCTGCGGCCCTCCTCCTGACATGCTATGCAATTC 105
Db      1 ATGAATCTGCTGATGCTCTCATGCTGCGGCCCTCCTCCTGACATGCTATGCAATTC 60
Qy      106 GGCCTGCAAACTCCTGAGAGCATGTTGAAAAGACATCAATTCGACATATCTATACCT 165
Db      61 GGCCTGCCCTTATTGAGAAATGTGATTTCCAGACAAATCAATCCACAAGTGTCTAAGACT 120
Qy      166 GAATACAAAGACCTTCTTCACAGAGTTCATAGACAGTATGCCGCTGCAGAGCTATGGGG 225
Db      121 GAATACAAAGACCTTCTTCACAGAGTTCATAGACAGCAATGCACTACAAATGCCATAGAT 180
Qy      226 AAATTCAAAGCACTGTTTCCTCCACAGCACTGACATGAACTGAAAACTTTGGACGATG 285
Db      181 GAATTGAAAGAACTGTTTCTTACCAAAAGCATGAAACTCTGAGCAATGTTGAGTGT 240
Qy      286 ATGCATACAGTGTACGACAGCATTTGGTAAATAT 320
Db      241 ATGCATTAATATATGACAGCAGTCTTGTGATTT 275
```

Search completed: January 25, 2003, 11:40:07
Job time : 89 secs


```
|||||
Db 1 ACAGCTGCCACGACGACTGACACACAGACGCGCCCTCCCATGAAAGCTGAGTG 60
QY 61 GTTCCTCATGCTGGCGCCCTCCTCCTGACTGCTATGACAGATTCTGGCTCAAACTCCG 120
Db 61 GTTCCTCATGCTGGCGCCCTCCTCCTGACTGCTATGACAGATTCTGGCTCAAACTCCG 120
QY 121 GAGGACATGTTGAAAAGACCATCAATTCGACATCTCTTACCTGAAATACAAAGGCTT 180
Db 121 GAGGACATGTTGAAAAGACCATCAATTCGACATCTCTTACCTGAAATACAAAGGCTT 180
QY 181 CTTCAAGATTCTATAGACAGTATGCGCTGCAGAGGCTATGCGGAAATTCAGCAGTGT 240
Db 181 CTTCAAGATTCTATAGACAGTATGCGCTGCAGAGGCTATGCGGAAATTCAGCAGTGT 240
QY 241 TTCCCTAACCACTGACATAGAACTCTGAAAACTTGGACTGATGATGATACAGTATAC 300
Db 241 TTCCCTAACCACTGACATAGAACTCTGAAAACTTGGACTGATGATGATACAGTATAC 300
QY 301 GACAGCATTGGTGTAAATATGACAGTAACTTACTTTACCCAAAGCGTTTGGCTCAGAG 360
Db 301 GACAGCATTGGTGTAAATATGACAGTAACTTACTTTACCCAAAGCGTTTGGCTCAGAG 360
QY 361 GCTACAGACTATGCGCAGAACTCATCTGTTGATTGCTAGAAAACCACTTCTCTTGTGT 420
Db 361 GCTACAGACTATGCGCAGAACTCATCTGTTGATTGCTAGAAAACCACTTCTCTTGTGT 420
QY 421 GCTTTTATGTGGGAAGCTGTAGACAACTGTTGAAACCTCAATTCATTCAATTCA 476
Db 421 GCTTTTATGTGGGAAGCTGTAGACAACTGTTGAAACCTCAATTCATTCAATTCA 476
```

RESULT 2

US-09-263-810-5
Sequence 5, Application US/09263810

Patent No. 6174992

GENERAL INFORMATION:

APPLICANT: Jian NI, Guo-Liang Yu and Reiner Gentz

TITLE OF INVENTION: Human Endometrial Specific Steroid-

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELIA, BYRNE, BAIN, GILFILLAN,

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/263,810

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/821,451

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-521 (PF257)

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-994-1700

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 476 BASE PAIRS

TYPE: NUCLEIC ACID

STRANDEDNESS: SINGLE

TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
US-09-263-810-5

Query Match 100.0%; Score 476; DB 4; Length 476;
Best Local Similarity 100.0%; Pred. No. 6,8e-142;

Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ACAGCTGCCACGACGACTGACACACAGACGCGCCCTCCCATGAAAGCTGAGTG 60
Db 1 ACAGCTGCCACGACGACTGACACACAGACGCGCCCTCCCATGAAAGCTGAGTG 60
QY 61 GTTCCTCATGCTGGCGCCCTCCTCCTGACTGCTATGACAGATTCTGGCTCAAACTCCG 120
Db 61 GTTCCTCATGCTGGCGCCCTCCTCCTGACTGCTATGACAGATTCTGGCTCAAACTCCG 120
QY 121 GAGGACATGTTGAAAAGACCATCAATTCGACATCTCTTACCTGAAATACAAAGGCTT 180
Db 121 GAGGACATGTTGAAAAGACCATCAATTCGACATCTCTTACCTGAAATACAAAGGCTT 180
QY 181 CTTCAAGATTCTATAGACAGTATGCGCTGCAGAGGCTATGCGGAAATTCAGCAGTGT 240
Db 181 CTTCAAGATTCTATAGACAGTATGCGCTGCAGAGGCTATGCGGAAATTCAGCAGTGT 240
QY 241 TTCCCTAACCACTGACATAGAACTCTGAAAACTTGGACTGATGATGATACAGTATAC 300
Db 241 TTCCCTAACCACTGACATAGAACTCTGAAAACTTGGACTGATGATGATACAGTATAC 300
QY 301 GACAGCATTGGTGTAAATATGACAGTAACTTACTTTACCCAAAGCGTTTGGCTCAGAG 360
Db 301 GACAGCATTGGTGTAAATATGACAGTAACTTACTTTACCCAAAGCGTTTGGCTCAGAG 360
QY 361 GCTACAGACTATGCGCAGAACTCATCTGTTGATTGCTAGAAAACCACTTCTCTTGTGT 420
Db 361 GCTACAGACTATGCGCAGAACTCATCTGTTGATTGCTAGAAAACCACTTCTCTTGTGT 420
QY 421 GCTTTTATGTGGGAAGCTGTAGACAACTGTTGAAACCTCAATTCATTCAATTCA 476
Db 421 GCTTTTATGTGGGAAGCTGTAGACAACTGTTGAAACCTCAATTCATTCAATTCA 476
```

RESULT 3

US-09-583-169-5
Sequence 5, Application US/09583169

Patent No. 6338948

GENERAL INFORMATION:

APPLICANT: Jian NI, Guo-Liang Yu and Reiner Gentz

TITLE OF INVENTION: Human Endometrial Specific Steroid-

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELIA, BYRNE, BAIN, GILFILLAN,

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/583,169

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/821,451

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.

REGISTRATION NUMBER: 33,073

:	REFERENCE/DOCKET NUMBER:	325800-521	(Pf257)
:	TELECOMMUNICATION INFORMATION:		
:	TELEPHONE:	201-994-1700	
:	TELEFAX:	201-994-1744	
:	INFORMATION FOR SEQ ID NO:	5:	
:	SEQUENCE CHARACTERISTICS:		
:	LENGTH:	476 BASE PAIRS	
:	TYPE:	NUCLEIC ACID	
:	STRANDEDNESS:	SINGLE	
:	TOPOLOGY:	LINEAR	
:	MOLECULE TYPE:	CDNA	
:	US-09-583-169-5		
<hr/>			
:	Query Match	100.0%;	Score 476; DB 4; Length 476;
:	Best Local Similarity	100.0%;	Pred. No. 6, 8e-142;
:	Matches 476; Conservative	0;	Mismatches 0; Indels 0; Gaps 0.
<hr/>			
OY	1	ACAGAGTCGCACCACGACTGAAGCAACAACAGCACCACCCCTCGCATTAAGAAGTGTGATG	60
Dd	1	ACAGAGTCGCCACGACGACTGAACAACAAGCACAGCACCACCCCTCGCATTAAGAAGTGTGATG	60
OY	61	GTCCTAATGCTGGCGGCCCTCTCTCGACATGCTATGCGAGATTGGCTGCGAACTCTTG	120
Dd	61	GTCCTAATGCTGGCGGCCCTCTCTCGACATGCTATGCGAGATTGGCTGCGAAACTCTTG	120
OY	121	GAGGACATGTTGTAANAAGACATCAATTCGACATATCTATACCTGGAATCAAAAGACTT	180
Dd	121	GAGGACATGTTGTAANAAGACATCAATTCGACATATCTATACCTGGAATCAAAAGACTT	180
OY	181	CTTCACAGATCTCATAGACAGTAGATCCGCTGACAGGCTATGGGAAAATTCACAGCATGT	240
Dd	181	CTTCACAGATCTCATAGACAGTAGATCCGCTGACAGGCTATGGGAAAATTCACAGCATGT	240
OY	241	TTCCTCACACGACGACATGAGACTCTGAAAAAATTGGACGATGATGATACAGGTATAC	300
Dd	241	TTCCTCACACGACGACATGAGACTCTGAAAAAATTGGACGATGATGATACAGGTATAC	300
OY	301	GACAGCATTTGGTGTATATATGAAAGTATTAATTAACCTTACCCCAAGCGTTTGGCTCAGAGG	360
Dd	301	GACAGCATTTGGTGTATATATGAAAGTATTAATTAACCTTACCCCAAGCGTTTGGCTCAGAGG	360
OY	361	GCTACAGACTATGGCCGCAAGCTATCTGTTGATTTGCTAGAAACACATTTCTTCTGTGTT	420
Dd	361	GCTACAGACTATGGCCGCAAGCTATCTGTTGATTTGCTAGAAACACATTTCTTCTGTGTT	420
OY	421	GCTTTTATGTGGGAACTGCTAGACAACTGTGAAACCTCAATTCAATTCATTCATTCA	476
Dd	421	GCTTTTATGTGGGAACTGCTAGACAACTGTGAAACCTCAATTCAATTCATTCATTCA	476
<hr/>			
:	RESULT 4		
:	US-08-969-987-5		
:	Sequence 5, Application US/08969987A		
:	Patent No. 6303297		
:	GENERAL INFORMATION:		
:	APPLICANT: Lincoln, Steve		
:	APPLICANT: Linger, Tod M.		
:	APPLICANT: Au-Young, Janice		
:	APPLICANT: Tang, Y. Tom I		
:	APPLICANT: Gould, Richard		
:	APPLICANT: Akerblom, Ingrid E.		
:	APPLICANT: Sellhammer, Jeffrey J.		
:	APPLICANT: Hawkins, Phillip R.		
:	APPLICANT: Murry, Lynn E.		
:	APPLICANT: Deleogene, Angelo M.		
:	APPLICANT: Levine, Wendy B.		
:	APPLICANT: Hillman, Jennifer L.		
:	APPLICANT: Goll, Surya K.		
:	APPLICANT: Altus, Christina M.		
:	APPLICANT: Labrie, Samuel T.		
:	APPLICANT: Shah, Purvi		
:	TITLE OF INVENTION: Database for Storage and Analysis of		

```

? TITLE OF INVENTION: Full Length Sequences
? FILE REFERENCE: 6514-069001
? CURRENT APPLICATION NUMBER: US/08/969,987A
?
? CURRENT FILING DATE: 1997-11-13
? EARLIER APPLICATION NUMBER: 08/282,955
?
? EARLIER FILING DATE: 1995-07-29
? EARLIER APPLICATION NUMBER: 08/187,530
?
? EARLIER FILING DATE: 1994-01-27
? EARLIER APPLICATION NUMBER: 08/179,873
?
? EARLIER FILING DATE: 1994-01-11
? EARLIER APPLICATION NUMBER: 08/100,523
?
? EARLIER FILING DATE: 1993-08-03
? EARLIER APPLICATION NUMBER: 07/977,780
?
? EARLIER FILING DATE: 1992-11-19
? EARLIER APPLICATION NUMBER: 07/916,491
?
? EARLIER FILING DATE: 1992-07-17
? EARLIER APPLICATION NUMBER: 08/289,822
?
? EARLIER FILING DATE: 1994-08-12
? EARLIER APPLICATION NUMBER: 08/581,240
?
? EARLIER FILING DATE: 1995-12-29
? EARLIER APPLICATION NUMBER: 08/657,697
?
? EARLIER FILING DATE: 1996-05-29
? EARLIER APPLICATION NUMBER: 08/747,547
?
? EARLIER FILING DATE: 1996-11-12
? EARLIER APPLICATION NUMBER: 08/712,710
?
? EARLIER FILING DATE: 1996-09-12
? EARLIER APPLICATION NUMBER: 08/744,026
?
? EARLIER FILING DATE: 1996-11-05
? EARLIER APPLICATION NUMBER: 08/786,999
?
? EARLIER FILING DATE: 1997-01-23
? EARLIER APPLICATION NUMBER: 08/822,262
?
? EARLIER FILING DATE: 1997-03-20
? EARLIER APPLICATION NUMBER: 08/951,750
?
? NUMBER OF SEQ ID NOS: 10
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 5
? LENGTH: 495
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-08-969-987-5

```

Query Match 47.0%; Score 223.8; DB 4; Length 495;
Best Local Similarity 71.4%; Pred. No. 1,1e-61;
Matches 325; Conservative 0; Mismatches 122; Indels 8; Gaps 2

OY 7 TGCACCGCAGCTGAACAACAGACAGACGCGCCCTGGCATTAAGCTGTGATGTCCTC 66
| | | | |
Db 7 TGCACCGCGAGTGAACACGACAGCACACCCTCACCATGAAGTTGATGTCTC 66
| | | | |
OY 67 ATGTGGCGGCCCTCTCCCTCGACATGCTATGCAAGTTTGCGCTGCAAACTCTGGAGAC 126
| | | | |
Db 67 ATGTGGCGGCCCTCTCCCGACGACTGTCTAGCGAGGCTGTGCTGCCCTTATTGGAAT 126
| | | | |
OY 127 ATGGTGAAGAAGCAATCAATTCGACATATCTATACCTGTAATAAAGAGCTTCTCAA 186
| | | | |
Db 127 GTGATTTCCAGCAATCAATCCACAGTGTCTAAGACTGAATACAAAGACTTCTCAA 186
| | | | |
OY 187 GAGTCTTAAGACAGTATGCGCGGTGAGAGGCTATGGGGAATTCAGCAGTGTCTCTC 246
| | | | |
Db 187 GAGTCTTAAGACCAATGCGCACTACCAAAATGCGATAGTAGTAATGAAAGATGTTTTCTT 246
| | | | |
OY 247 AACCACTACATAGAACTGTGA AAAA CTTTGGACTGATGATGATACATACAGTTCACGACG 306
| | | | |
Db 247 AACCAAAAGGATAAAACCTGTGAGCAATGTTGAGGTATTATGCAATTAAATATGACAGC 306
| | | | |
OY 307 ATTGGGTATATGAAGAGTAAATTAACTTTACCCAAGCGCTTTGGCTGAGAGGCTACA 366
| | | | |
Db 307 AGCTTTTGTGATTT-----ATTTAACCTTTCTGCAAAACCTTTGGCTCACAGAACTGCA 360
| | | | |
OY 367 GACTATGCCAGAACATCATCTGTGATTGTGCTGGAAC--CACTTTCTTTGTGTCTT 424
| | | | |
Db 361 GGATAGGTAGAAACCAACTACTCGAGTCTCTGCAAAACAGCACTTCTTCTTATATCTC 420
| | | | |

Qy 425 TTTATGTCGGAACTGCTAGACAACGTGTTGAAACCT 459
||| | ||| | ||||| ||||| |||||
Db 421 TTTTACTACAAACTACAAGACAATTGTTGAAACCT 455

RESULT 5

```

: US-08-455-896-1
: Sequence 1, Application US/08455896
: Patent No. 5668267
:
: GENERAL INFORMATION:
:
: APPLICANT: MATSON, MARK A.
: APPLICANT: FLEMING, TIMOTHY P.
: TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
: TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ROGERS, HOWELL & HAFERKAMP
: STREET: 7733 FORSTH BOULEVARD, SUITE 1400
: CITY: ST. LOUIS
: STATE: MISSOURI
: COUNTRY: USA
: ZIP: 63105-1817
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/455,896
: FILING DATE:
:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: HOLLAND, DONALD R.
: REGISTRATION NUMBER: 35,197
: REFERENCE/DOCKET NUMBER: 952726
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (314) 727-5188
: TELEFAX: (314) 727-6092
: INFORMATION FOR SEQ ID NO: 1:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 503 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHEICAL: NO
:
: ANTI-SENSE: NO
:
: US-08-455-896-1

```

Query Match	47.08;	Score 223.8;	DB 1;	Length 503;
Best Local Similarity	71.48;	Pred. No. 1.1e-61;		
Matches 325; Conservative	0;	Mismatches 122;	Indels 8;	Gaps 2

0Y	7	TCCACGACGAGCTAATACACAGACAGACGGCGCTGGCCATGAGCTGCTATGGTCTC	66
Db	22	TGCGACCCCGAGCTAATACACGACAGCAGCAGCTCAACCATGAGTGGTGGTGGCTC	81
0Y	67	ATCGTGGGGGGCCCTCTCTGCTACCTGGTATGACGATTTCTGGCTGCAACCTCTGGAGAC	128
Db	82	ATGCTGGGGGGCCCTCTGCCAGACCTGCTAGCAGAGCTCTGGCTGCCCTCTATGGACAT	144
0Y	127	ATGCTGGAATAACCATCAATATTCGACATCTCTATACCTGATACAAAGAGCTTTCGA	186
Db	142	GTTATTTCCAAACATATACATATCCACAGTGTCTTAGACATGAAATACAAAGAAATCTTTGAA	201
0Y	187	GACTTATAGACATGATATGCGCGTGCAGAGGCTATGGGGAAATTCAGCAGGTTCCTC	248
Db	202	GAGTTTATAGACACAAATGGCCATTCACAAATGGCCATATGATTAATGAGGAATGTTTTCTT	261
0Y	247	AACCACTGCATAGAACTTGTAATAAATCTTTGCACTGATGATGCATACAGTGTAGACAGC	306
Db	262	AACCAAAACGATGAACACTGTAGGCAAAATGTGTGAGTGTATATCAATTAATATATAGACAGC	321

Oy	307	ATTGGGTGTAATATCAAGAAATAATTACCTTTACCACAGGCGTTGGCTCAGAGGGCTTCA	366
Db	322	AGCTCTTGTGATT-----ATTTACCTTTCTGCAAGACCTTTGGCTCACAGAACTGCA	375
Oy	367	GACTATGGCCACAACCTCATCTGTGATTTGGTATAAAC--CACTTTCTCTGTGTGCTT	424
Db	376	GGGTATGGTGAATAAACCAACTACGAGATTGCTGCAACCACTCTCTCTTATATGCT	435
Oy	425	TTTATGTGGGAACCTCTGACACACTGTGAAACCT	459
Db	436	TTTTACTCAAACTCAAGACAAATTTGTTGAAACCT	470

RESULT 6

```

US-08-933-149-1
: Sequence 1, Application US/08933149
: Patent No. 5922836
:
: GENERAL INFORMATION:
: APPLICANT: MATSON, MARK A.
: APPLICANT: FLEMING, TIMOTHY P.
: TITLE OF INVENTION: MAMMOGLOBIN, A SECRETED
: TITLE OF INVENTION: MAMMARY SPECIFIC BREAST CANCER PROTEIN
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: HOWELL & HAPERKAMP, L.C.
: STREET: 773 FORSTH BOULEVARD, SUITE 1400
: CITY: ST. LOUIS
: STATE: MISSOURI
: COUNTRY: USA
: ZIP: 63105-1817
: COMPUTER READABLE FORM:
: MEDIUM TYPE: FLOPPY disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/933,149
: FILING DATE:
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: HENDERSON, MELDIE W.
: REGISTRATION NUMBER: 37,848
: REFERENCE/DOCKET NUMBER: 6029-6040
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (314) 727-5188
: TELEFAX: (314) 727-6092
: INFORMATION FOR SEQ. ID NO. 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 503 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOHETICAL: NO
: ANTI-SENSE: NO
:
: US-08-933-149-1

```

Query Match	47.08;	Score 223.8;	DB 2;	Length 503;
Best Local Similarity	71.48;	Pred. No. 1.1e-61;		
Matches 325; Conservative	0;	Mismatches 122;	Indels 8;	Gaps 2

Qy	7	TGCACGACGACGATGAACACAGAGAGAGCGCGCTGCCATGAACCTCATATGGTCTC	66
Db	22	TGCACCGCGCGATGAACACGAGCAGCAGCTCACCATGAATTCGTATGGTCTC	81
Qy	67	ATGCTGGGCGCCCTCTCTCTCACTGCTATGACGATTCCTGGCTGGCAACCTCTCGAAGAC	128
Db	82	ATGCTGGCGCCCTCTCTCCACACCTGGTATGACAGGCTCTGGCTGGCCCTTATGGACAAAT	141
Qy	127	ATGGTTGAAGACCAATCAATTCGACATATCTTAACTGAAATACAAAGACCTTCTTCAA	186
Db	142	GTGATTTCCAGACCAATATCAATCCACAGTGCTTCAAGACTGAAATACAAAGACTCTTCAA	201


```

:   LENGTH: 535
:   TYPE: DNA
:   ORGANISM: Homo Sapiens
US-09-215-818-1

Query Match
Best Local Similarity 47.0%; Score 223.8; DB 4; Length 535;
Best Local Similarity 71.4%; Pred. No. 1.2e-61;
Matches 325; Conservative 0; Mismatches 122; Indels 8; Gaps 2;

OY 7 TGGCAGCGACGACTGAAACAGACAGACGCGCGCTCCGCTGGAAGTGGTGAATGGTCCTC 66
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 43 TGGCAGCGCGCGACTGAACACGACAGACGACGCGCTCCGCTGGAAGTGGTGAATGGTCCTC 102
OY 67 ATGCTGCGCGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 126
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 103 ATGCTGCGCGCGCTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 162
OY 127 ATGGTTGAAAGACCATCAATTCGACATATCTATACCTGATATCAAGAGAGCTTCTTCAA 186
    || || ||||| ||||| || || || || || || || || || || || || || || || || ||
DB 163 GTGATTTCCAAACATCAATTCACAAAGTGTCTAAGCTGAATACAAAGAACTTCTTCAA 222
OY 187 GAGTTGATGACAGTATGCGGCTGACAGAGCTATGGGAAATTCAGACAGTGTCTCTC 246
    ||||| ||||| ||||| || || || || || || || || || || || || || || || || ||
DB 223 GAGTTGATGACAGTATGCGGCTGACAGAGCTATGGGAAATTCAGACAGTGTCTCTC 282
OY 247 AACCACTGACATAGAGTCTGAAAACTTTGGAGTATGATGATACAGTATGACAGCAGC 306
    ||||| || || ||||| || || || || || || || || || || || || || || || || ||
DB 283 AACCAAGCGATGAACCTCTGACATCTGAGTGTATGCAATTAATATATGACAGC 342
OY 307 ATTGGTGAATATGAGAGTATTAATTACCTTACCAAGCGGTGGCTGACAGCGCTTACA 366
    || || ||||| || || ||||| || || ||||| || || ||||| || || || || || || ||
DB 343 AGCTTTGATGATTT-----ATTTAACTTTCTGCAAGACCTTGGCTACAGAACTGCA 396
OY 367 GACTATGCGCAGACTCATCTGTTGATGCTAGAAAC--CAGTTTCTCTGTTGTTGCT 424
    || ||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 397 GGATATGCGTGAACCAACGATGCGATGCTGCAAAACCAACCTTCTTCTTATGTCT 456
OY 425 TTTATGTGGAACTGCTGACAGACTTTGAAACCT 459
    || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 457 TTTACTACAACTACAGACAAATTTGTAACCT 491

RESULT 11
US-08-455-896-5
: Sequence 5, Application US/08455896
: Patent No. 5688267
: GENERAL INFORMATION:
: APPLICANT: WATSON, MARK A.
: APPLICANT: FLEMING, TIMOTHY P.
: TITLE OF INVENTION: DNA SEQUENCE AND ENCODED
: TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN
: NUMBER OF SEQUENCES: 13
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: ROGERS, HOWELL & HAFERKAMP
: STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
: CITY: ST. LOUIS
: STATE: MISSOURI
: COUNTRY: USA
: ZIP: 63105-1817
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/455,896
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: HOLLAND, DONALD R.
: REGISTRATION NUMBER: 35,197
: REFERENCE/DOCKET NUMBER: 952726
: TELECOMMUNICATION INFORMATION:
```

```

:   TELEPHONE: (314) 727-5188
:   TELEFAX: (314) 727-6092
:   INFORMATION FOR SEQ ID NO: 5:
:   SEQUENCE CHARACTERISTICS:
:   LENGTH: 403 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:   MOLECULE TYPE: cDNA to mRNA
:   HYPOTHETICAL: NO
:   ANTI-SENSE: NO
US-08-455-896-5

Query Match
Best Local Similarity 42.3%; Score 201.2; DB 1; Length 403;
Best Local Similarity 71.9%; Pred. No. 1.6e-54;
Matches 279; Conservative 0; Mismatches 103; Indels 6; Gaps 1;

OY 7 TGGCAGCGACGACTGAAACAGACAGACGCGCGCTCCGCTGGAAGTGGTGAATGGTCCTC 66
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 22 TGGCAGCGCGCGACTGAACACGACAGCAGCGCTCCGCTGGAAGTGGTGAATGGTCCTC 81
OY 67 ATGCTGCGCGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 126
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 82 ATGCTGCGCGCGCTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 141
OY 127 ATGGTTGAAAGACCATCAATTCGACATATCTATACCTGATATCAAGAGAGCTTCTTCAA 186
    || || ||||| ||||| || || || || || || || || || || || || || || || || ||
DB 142 GTGATTTCCAAACATCAATTCACAAAGTGTCTAAGCTGAATACAAAGAACTTCTTCAA 201
OY 187 GAGTTGATGACAGTATGCGGCTGACAGAGCTATGGGAAATTCAGACAGTGTCTCTC 246
    ||||| ||||| ||||| || || || || || || || || || || || || || || || || ||
DB 202 GAGTTGATGACAGTATGCGGCTGACAGAGCTATGGGAAATTCAGACAGTGTCTCTC 261
OY 247 AACCACTGACATAGAGTCTGAAAACTTTGGAGTATGATGATACAGTATGACAGCAGC 306
    ||||| ||||| ||||| || || || || || || || || || || || || || || || || ||
DB 262 AACCAAGCGATGAACCTCTGACATCTGAGTGTATGCAATTAATATATGACAGC 321
OY 307 ATTGGTGAATATGAGAGTATTAATTACCTTACCAAGCGGTGGCTGACAGCGCTTACA 366
    || || ||||| || || ||||| || || ||||| || || ||||| || || || || || || ||
DB 322 AGCTTTGATGATTT-----ATTTAACTTTCTGCAAGACCTTGGCTACAGAACTGCA 375
OY 367 GACTATGCGCAGACTCATCTGTTGATGCTAGAAAC--CAGTTTCTCTGTTGTTGCT 394
    || ||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 376 GGATATGCGTGAACCAACGATGCGATT 403

RESULT 12
US-08-933-149-5
: Sequence 5, Application US/08933149
: Patent No. 5922836
: GENERAL INFORMATION:
: APPLICANT: WATSON, MARK A.
: APPLICANT: FLEMING, TIMOTHY P.
: TITLE OF INVENTION: MAMMAGLOBIN, A SECRETED
: TITLE OF INVENTION: MAMMARY SPECIFIC BREAST CANCER PROTEIN
: NUMBER OF SEQUENCES: 14
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: HOWELL & HAFERKAMP, L.C.
: STREET: 7733 FORSYTH BOULEVARD, SUITE 1400
: CITY: ST. LOUIS
: STATE: MISSOURI
: COUNTRY: USA
: ZIP: 63105-1817
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/933,149
: FILING DATE:
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
```



```

? ZIP: 63105-1817
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC Compatible
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/9/082,253
? FILING DATE:
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/455,896
? FILING DATE: 05/31/1995
? ATTORNEY/AGENT INFORMATION:
? NAME: HOLLAND, DONALD R.
? REGISTRATION NUMBER: 35,197
? REFERENCE/DOCKET NUMBER: 952726
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (314) 727-5188
? TELEFAX: (314) 727-6092
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 403 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA to mRNA
? HYPOTHEICAL: NO
? ANTI-SENSE: NO
? US-09-082-253-5

```

Query Match 42.3%; Score 201.2; DB 3; Length 403;

Best Local Similarity 71.9%; Pred. No. 1.6e-54;

Matches 279; Conservative 0; Mismatches 103; Indels 6; Gaps 1;

```

OY 7 TGCACGACGACGTGACACAGACAGCGCGCTGCCATGAAGCTGTGATGCTCTC 66
    |||||
DB 22 TGCACCGCGCGCTGACACAGCGCGCTGCCATGAAGCTGTGATGCTCTC 81
OY 67 ATGCTGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 126
    |||||
DB 82 ATGCTGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 141
OY 127 ATGCTGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 186
    |||||
DB 142 ATGCTGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 201
OY 187 GAGTTCATGACAGTGTGCGCTGACAGGCTATGCGGAATTCAGAGCTGTTCTC 246
    |||||
DB 202 GAGTTCATGACAGTGTGCGCTGACAGGCTATGCGGAATTCAGAGCTGTTCTC 261
OY 247 AACCACTGACATGACATCTGAAAACTTTGGAGTGTGATGATGATGAGGATG 306
    |||||
DB 262 AACCACTGACATGACATCTGAAAACTTTGGAGTGTGATGATGATGAGGATG 321
OY 307 ATTTGGTATATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 366
    |||||
DB 322 AGCTTTGATGATTT-----ATTTAACTTTCTGCAAGACCTTGGCTCAGAACTGCA 375
OY 367 GACTATGCCAGACTCATCTGTGATT 394
    |||||
DB 376 GGTATGTGAGAAACCACTACGATT 403

```

RESULT 15

PCT-US96-08235-5

Sequence 5, Application PC/TUS9608235

GENERAL INFORMATION:

APPLICANT: WATSON, MARK A.

APPLICANT: FLEMING, TIMOTHY P.

TITLE OF INVENTION: DNA SEQUENCE AND ENCODED

TITLE OF INVENTION: MAMMARY-SPECIFIC BREAST CANCER PROTEIN

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: ROGERS, HOWELL & HAFERKAMP

STREET: 7733 FORSYTH BOULEVARD, SUITE 1400

CITY: ST. LOUIS

STATE: MISSOURI

COUNTRY: USA

ZIP: 63105-1817

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC Compatible

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/08235

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: HOLLAND, DONALD R.

REGISTRATION NUMBER: 35,197

REFERENCE/DOCKET NUMBER: 964796

TELECOMMUNICATION INFORMATION:

TELEPHONE: (314) 727-5188

TELEFAX: (314) 727-6092

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 403 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

HYPOTHEICAL: NO

ANTI-SENSE: NO

PCT-US96-08235-5

Query Match 42.3%; Score 201.2; DB 5; Length 403;

Best Local Similarity 71.9%; Pred. No. 1.6e-54;

Matches 279; Conservative 0; Mismatches 103; Indels 6; Gaps 1;

```

OY 7 TGCACGACGACGTGACACAGACAGCGCGCTGCCATGAAGCTGTGATGCTCTC 66
    |||||
DB 22 TGCACCGCGCGCTGACACAGCGCGCTGCCATGAAGCTGTGATGCTCTC 81
OY 67 ATGCTGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 126
    |||||
DB 82 ATGCTGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 141
OY 127 ATGCTGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 186
    |||||
DB 142 ATGCTGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 201
OY 187 GAGTTCATGACAGTGTGCGCTGACAGGCTATGCGGAATTCAGAGCTGTTCTC 246
    |||||
DB 202 GAGTTCATGACAGTGTGCGCTGACAGGCTATGCGGAATTCAGAGCTGTTCTC 261
OY 247 AACCACTGACATGACATCTGAAAACTTTGGAGTGTGATGATGATGAGGATG 306
    |||||
DB 262 AACCACTGACATGACATCTGAAAACTTTGGAGTGTGATGATGATGAGGATG 321
OY 307 ATTTGGTATATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 366
    |||||
DB 322 AGCTTTGATGATTT-----ATTTAACTTTCTGCAAGACCTTGGCTCAGAACTGCA 375
OY 367 GACTATGCCAGACTCATCTGTGATT 394
    |||||
DB 376 GGTATGTGAGAAACCACTACGATT 403

```

Search completed: January 25, 2003, 10:29:14

Job time : 76 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 25, 2003, 11:33:34 : Search time 2256 Seconds
(without alignments)
3417.133 Million cell updates/sec

Title: US-09-806-302A-1
476
Perfect score: 1 acagcgtccacgcacgact.....ccctcaatcattcattca 476
Sequence:

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estnu:*
4: em_estnu:*
5: em_estnu:*
6: em_estnu:*
7: em_estnu:*
8: em_estnu:*
9: em_estnu:*
10: em_estnu:*
11: em_estnu:*
12: em_estnu:*
13: em_estnu:*
14: em_estnu:*
15: em_estnu:*
16: em_estnu:*
17: em_estnu:*
18: em_estnu:*
19: em_estnu:*
20: em_estnu:*
21: em_estnu:*
22: em_estnu:*
23: em_estnu:*
24: em_estnu:*
25: em_estnu:*
26: em_estnu:*
27: em_estnu:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	429.2	50.2	636	10	AW966513 EST378587
2	428.2	50.2	647	10	AW966509 EST378583
3	408.2	85.8	494	10	AA451131 UI-H-B13
4	407.4	85.6	491	9	AA393164 zt73f02.x
5	405.2	85.1	434	13	BI964837 t167c01.y
6	397.4	83.5	464	9	AI800231 t176c07.x

C	7	391.6	82.3	460	9	AI659370
C	8	389.6	81.8	471	9	AI936084
C	9	389	81.7	458	10	AW779377
C	10	385.2	80.9	458	9	AI491987
C	11	378.2	79.5	396	9	AA297402
C	12	366.6	77.0	435	9	AA398560
C	13	364.6	76.6	435	9	AA493295
C	14	357	75.0	365	10	BE044895
C	15	343	72.1	351	10	BE044893
C	16	341.2	71.7	418	10	BE044893
C	17	338	71.0	357	10	AW510561
C	18	335.4	70.5	456	10	AW966498
C	19	294	61.8	378	13	BI963186
C	20	280.2	58.9	318	9	AA298750
C	21	280	58.8	368	9	AA525178
C	22	249.6	52.4	268	9	AA297452
C	23	247	51.9	253	9	AA297456
C	24	240	50.4	242	9	AA297405
C	25	206.4	43.4	463	9	AI791322
C	26	191.8	40.3	510	10	AW972474
C	27	190.2	40.0	388	10	AW175618
C	28	188	39.5	447	9	AA513640
C	29	186.8	39.2	372	9	AI904282
C	30	186.2	39.1	375	9	AI904278
C	31	185	38.9	189	9	AA297482
C	32	184.8	38.8	444	9	AA931451
C	33	181.4	38.1	436	9	AI732534
C	34	169	35.5	421	9	AI124690
C	35	165.4	34.7	184	9	AI937042
C	36	160	33.6	197	12	BE745480
C	37	157	33.0	309	14	BO304242
C	38	155.2	32.6	294	9	AI905273
C	39	143	30.0	263	9	AI905265
C	40	140.6	29.5	392	9	AA513647
C	41	113	23.7	282	10	AW175603
C	42	112.2	23.6	127	12	BG219371
C	43	110.6	23.2	193	10	AW207696
C	44	104.4	21.9	511	13	BI278959
C	45	104	21.8	510	13	BI285288

ALIGNMENTS

RESULT 1
LOCUS AW966513 636 bp mRNA linear EST 01-JUN-2000
DEFINITION EST378587 MAGE resequences, MAGE Homo sapiens CDNA, mRNA sequence.
ACCESSION AW966513
VERSION AW966513.1 GI:8156349
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 636)
Hedge, P., Qi, R., Abernathy, K., Dharp, S., Caspard, R., Gay, C., Holt,
, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J., and
Quackenbush, J.
Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
unpublished (2000)
JOURNAL Contact: John Quackenbush
COMMENT The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
plate: 229
Seq primer: Forward.
Location/Qualifiers
1..636
/organism="Homo sapiens"


```

/db.xref="taxon:9606"
/clone_id="NCL_CGAP_Sub5"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; NCL_CGAP_Sub5
is a subtracted library derived from NCL_CGAP_Sub4. The
NCL_CGAP_Sub5 library had 3 million recombinants. A
single-stranded DNA preparation of NCL_CGAP_Sub5 was used
as a tracer in a subtractive hybridization with a driver
comprising: the IMAGE pool (NCL_CGAP_Kid3 pool 1 LLAM
3334-3337, 3682-3683, 3798-3803) (IMAGE clones:
1323376-1323911, 1456008-1456775, 1500552-1502855);
NCL_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778
(IMAGE clones: 1323912-1325831, 1471368-1472903,
1492104-1493255); NCL_CGAP_Lu5 pool 1 LLAM 3575-3582,
3851-3854 (IMAGE clones: 1414920-1417991, 1520904-1522429
); NCL_CGAP_Gc4 pool 1 LLAM 3164-3167, 3716-3720,
3733-3735 (IMAGE clones: 1257096-1258631, 1469064-1470983,
1475592-1476743); NCL_CGAP_Pf22 pool 1 LLAM 2457-2459,
2758-2759, 3062-3068 (IMAGE clones: 985608-986759
, 1101192-1101959, 1217928-1220615); NCL_CGAP_Co10 pool 1
LLAM 2644-2653, 2871-2872 (IMAGE clones: 1057416-1061255
, 114584-1145351). (10% of the driver population), plus a
pool of 3,840 arrayed clones from NCL_CGAP_Sub1 (IMAGE
clones: 2708616-2710535) and NCL_CGAP_Sub2 (IMAGE
clones: 2710536-2712455) (10% of the driver population
), plus a pool of 11,136 clones from NCL_CGAP_Sub3 (IMAGE
clones: 2712456-2723591) (10% of the driver population),
plus a pool of 5,472 clones from NCL_CGAP_Sub4 (IMAGE
clones: 2723592-2728969) (70% of the driver population).
Subtraction was performed as previously described [Bonaldo
, Lennon & Soares (1996)]: Normalization and Subtraction:
Two Approaches To Facilitate Gene Discovery. Genome
Research 6, 791-806.
TAG_Lib=NCL_CGAP_Br2
TAG_Tissue=breast
TAG_SEQ=AAACC"
BASE COUNT      132 a      98 c      109 g      154 t      1 others
ORIGIN

```

```

Query Match      85.8% Score 408.2; DB 10; Length 494;
Best Local Similarity 96.4%; Pred. No. 3.7e-106;
Matches 429; Conservative 0; Mismatches 13; Indels 3; Gaps 1;

OY 23 ACACAGACAGAGCCGCTCGGCATGAAGCTGCTGATGCTCATGCTGGCGCCCTCC 82
|||||
DB 494 ACACAGACAGAGCCGCTCGGCATGAAGCTGCTGATGCTCATGCTGGCGCCCTCC 435
|||||

OY 83 TCCTGACGCTGATGACATGCTGCTGCTGCAACCTCCGAGACATGCTTGAAGACCA 142
|||||
DB 434 TCATGACGCTGATGACATGCTGCTGCTGCAACCTCCGAGACATGCTTGAAGACCA 375
|||||

OY 143 TCAATTCGACATATCTATACCTGTAATCAAGAAGCTTCTCAAGATTCATAGACAGT 202
|||||
DB 374 TCAATTCGACATATCTATACCTGTAATCAAGAAGCTTCTCAAGATTCATAGACAGT 315
|||||

OY 203 ATGCGCGCTGACAGGCTATGGGAAATCAAGCAGCTTTCCTCAACAGTCACATAGAA 262
|||||
DB 314 ATGCGCGCTGACAGGCTATGGGAAATCAAGCAGCTTTCCTCAACAGTCACATAGAA 255
|||||

OY 263 CTCGTGAAGAACTTTGGAGTGAATGATGATACAGTATGACAGCATTGGTGAATATGA 322
|||||
DB 254 CTCGTGAAGAACTTTGGAGTGAATGATGATACAGTATGACAGCATTGGTGAATATGA 195
|||||

OY 323 AGAGTAATTAATTAACCAAGCGCTTGGCTCAGAGGCTACAGACTATAGGCACAAAT 382
|||||
DB 194 AGAGTAATTAATTAACCAAGCGCTTGGCTCAGAGGCTACAGACTATAGGCACAAAT 135
|||||

OY 383 CATCTGTTATGCTAGAAACACATCTTCTTGTG---TTGCTTTTATGGAAGCT 439
|||||
DB 134 CATCTGTTATGCTAGAAACACATCTTCTTGTGTTGTGCTTTTATGGAAGCT 75
|||||

```

```

OY 440 CTAGACACTGTTGAACCTCAATT 464
|||||
DB 74 CTAGACACTGTTGAACCTCAAAAT 50
|||||

RESULT 4
AA393164
LOCUS
DEFINITION
AA393164 491 bp mRNA linear EST 16-MAY-1997
2173702.r1 Soares-testis_NHT Homo sapiens cDNA clone IMAGE:727995
5' similar to TR:G1199596 G1199596 MAMMALOBLIN.; mRNA sequence.
AA393164
VERSION
AA393164.1 GI:2046134
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 491)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B.,
Schellenberg, R., Stepien, M., Tan, F., Thelings, B., White, Y., Wyllie,
T., Waterston, R. and Wilson, R.
WashU-Merck EST Project 1997
JOURNAL
Unpublished (1997)
COMMENT
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 380.
Location/Qualifiers
1. 491
/organism="Homo sapiens"
/db.xref="GDB:5924907"
/db.xref="taxon:9606"
/clone="IMAGE:727995"
/clone_lib="Soares-testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      149 a      109 c      100 g      133 t
ORIGIN
Query Match      85.6% Score 407.4; DB 9; Length 491;
Best Local Similarity 97.0%; Pred. No. 6.2e-106;
Matches 426; Conservative 0; Mismatches 11; Indels 2; Gaps 1;

OY 28 GACAGAGCGCGCTCGCCATGAAGCTGCTGATGCTCTCAATGCTGGCGCCCTCTCTCG 87
|||||
DB 1 GACAGAGCGCGCTCGCCATGAAGCTGCTGATGCTCTCAATGCTGGCGCCCTCTCTCG 60
|||||

OY 88 CACTGCTATGACGATTTGGCTGCAAACTCTCGAGACATGGTTGAAGACATCAAT 147
|||||
DB 61 CACTGCTATGACGATTTGGCTGCAAACTCTCGAGACATGGTTGAAGACATCAAT 120
|||||

OY 148 TCCGACATATCTATACCTGAATATCAAGAGCTTCTTCAAGATTCTATAGACAGTATGCC 207
|||||
DB 121 TCCGACATATCTATACCTGAATATCAAGAGCTTCTTCAAGATTCTATAGACAGTATGCC 180
|||||

OY 208 GCTGCGAGAGGCTATGGGAAATTCAGCAGTGTTCCTCAACAGTCACATAGAACTCTG 267
|||||

```


found through the I.M.A.G.E. Consortium/ULNL at:

www.bio.lnl.gov/db/ftp/image/image.html

Insert Length: 575 Std Error: 0.00

Seq primer: -400P from Gibco

High quality sequence stop: 453

Location/Qualifiers

1. .464

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2137932"

/clone_lib="NCI_CGAP_Kid11"

/lab_host="DH10B"

/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI;

Plasmid DNA from the normalized library NCI_CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP

purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs

from a pool of 5,000 clones made from the same library (cloneids 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M.

Fatima Bonaldo.

BASE COUNT 131 a 97 c 105 g 131 t

ORIGIN

Query Match 83.5%; Score 397.4; DB 9; Length 464;

Best Local Similarity 96.8%; Pred. No. 4,5e-103;

Matches 417; Conservative 1; Mismatches 11; Indels 3; Gaps 1;

37 CGCCCTCGCATGAGAGCTGCTGATGATGCTCATCTGCGCGCCCTCCCTGCTGCTAT 96

464 CGCCCTCGCATGAGAGCTGCTGATGATGCTCATCTGCGCGCCCTCCCTGCTGCTAT 405

97 GCAAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 156

404 GCAAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 345

157 TCTATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 216

344 TCTATACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 285

217 GCTATGAGGAATTCAGACAGTGTCTTCTCAACAGTCATGATGATGATGATGATG 276

284 GCTATGAGGAATTCAGACAGTGTCTTCTCAACAGTCATGATGATGATGATGATG 225

277 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 336

224 GACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 165

337 TACCAAGGCGTTGGCTCAGAGGCTACAGACTATGCGCAACATCTGTTGATTCG 396

164 TACCAAGGCGTTGGCTCAGAGGCTACAGAGTATGCGCAACATCTGTTGATTCG 105

397 TAGAAGCACTTCTCTCTG--TTGCTTTTATGCGAAGTCTGAGCAACTGTTG 453

104 TAGAAGCACTTCTCTCTG--TTGCTTTTATGCGAAGTCTGAGCAACTGTTG 45

454 AACCTCAAT 464

44 AACCTCAAT 34

RESULT 7

AI659370/c 460 bp mRNA linear EST 15-DEC-1999

LOCUS tui1c09.x1 NCI_CGAP Pr28 Homo sapiens cDNA clone IMAGE:2250736 3'

Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:

Mammalia: Eutheria: Primates: Catarrhini: Hominiidae: Homo.

1 (bases 1 to 460)

NCI_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgaps@emall.nih.gov

Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

clone distribution: NCI_CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/ULNL at:

www.bio.lnl.gov/db/ftp/image/image.html

Insert Length: 559 Std Error: 0.00

Seq primer: -400P from Gibco

High quality sequence stop: 448.

Location/Qualifiers

1. .460

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2250736"

/clone_lib="NCI_CGAP_Pr28"

/sex="male"

/dev_stage="adult"

/lab_host="DH10B"

/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the

normalized library NCI_CGAP_Pr28 was prepared, and ss

circles were made in vitro. Following HAP purification,

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (cloneids

985608-986759, 1101192-1101959, and 1217928-1220615).

Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 132 a 94 c 101 g 133 t

ORIGIN

Query Match 82.3%; Score 391.6; DB 9; Length 460;

Best Local Similarity 97.2%; Pred. No. 2e-101;

Matches 410; Conservative 9; Mismatches 9; Indels 3; Gaps 1;

46 ATGAAGCTGCTGATGATGCTGATGATGCTGCGCCCTCTCTGCTGCTATGATGATCT 105

460 ATGAAGCTGCTGATGATGCTGATGATGCTGCGCCCTCTCTGCTGCTATGATGATCT 401

106 GCGTCAAACTCTGAGAGACATGTTGAAAGACCATCAATTCCGACATATCTATACCT 165

400 GCGTCAAACTCTGAGAGACATGTTGAAAGACCATCAATTCCGACATATCTATACCT 341

166 GATACCAAGAGCTTCTTCAAGAGTTCATGACAGTATGCGCGTGCAGAGGCTATGGG 225

340 GATACCAAGAGCTTCTTCAAGAGTTCATGACAGTATGCGCGTGCAGAGGCTATGGG 281

226 AATATCAAGAGTGTCTTCTCAACCACTGACATAGAACTGTGAAAACCTTGGAGTATG 285

280 AATATCAAGAGTGTCTTCTCAACCACTGACATAGAACTGTGAAAACCTTGGAGTATG 221

286 ATGATACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 345

220 ATGATACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 161

346 GCTTTGGCTCAGAGGCTACAGACTATGCGCAGAACTCATCTGTTGATTTGCTAGAACCA 405

160 GCTTTGGCTCAGAGGCTACAGACTATGCGCAGAACTCATCTGTTGATTTGCTAGAACCA 101

406 CTTTCTCTTGTG--TTGCTTTTATGCGAAGTCTGAGCAACTGTTGAAACCTCAA 462

100 CTTTCTCTTGTG--TTGCTTTTATGCGAAGTCTGAGCAACTGTTGAAACCTCAA 41

RESULT 8	AI936084	471 bp	MRNA	linear	EST 02-SEP-1999
LOCUS	AI936084/c				
DEFINITION	W61d06.x1 NCI-CGAP_P122 Homo sapiens CDNA clone IMAGE:2459819 3'				
	similar to TR:075556 075556 MAMMAGLOBIN B PRECURSOR. ;, mRNA				
SEQUENCE.					
AI936084					
VERSION	AI936084.1	GI:5674954			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 471) 1				
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.				
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),				
JOURNAL	Tumor Gene Index				
COMMENT	Unpublished (1997)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgapbs-remail.nih.gov				
	Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.				
	Emmert-Buck, M.D., Ph.D.				
	CDNA Library Preparation: M. Bento Soares, Ph.D.				
	DNA Library Arrayed by: Greg Lennon, Ph.D.				
	DNA Sequencing by: Washington University Genome Sequencing Center				
	clone distribution: NCI-CGAP clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LLNL at:				
	www.dio.llnl.gov/dbfp/image/image.html				
	Seq primer: 40UP from Gibco				
	High quality sequence stop: 433.				
FEATURES	Location/Qualifiers				
SOURCE	1..471				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:2459819"				
	/clone_1lib="NCI-CGAP_P122"				
	/sex="male"				
	/tissue_type="normal prostate"				
	/lab_host="DH10B"				
	/note="Organ: prostate; Vector: pT73D-Pac (Pharmacia)				
	with a modified polylinker: 1st strand cDNA was prepared				
	from normal prostate bulk tissue, and was then primed with				
	a Not I - oligo(dT) primer. Double-stranded cDNA was				
	ligated to Eco RI adaptors (Pharmacia), digested with Not				
	I and cloned into the Not I and Eco RI sites of the				
	modified pT73 vector. Library is normalized, and was				
	constructed by Bento Soares and M. Fatima Bonaldo."				
BASE COUNT	131 a 94 c 103 g 143 t				
ORIGIN					
Query Match	81.8% Score 389.6; DB 9; Length 471;				
Best Local Similarity	97.1%: Pred. No. 7.7e-101;				
Matches 408; Conservative	0; Mismatches 9; Indels 3; Gaps 1.				
46	ATGACCTGCTGATGCTCATGCTGCGGCGCTCTCGACACTGCTATGAGATTC 105				
471	ATGAGCTGCTGATGCTCATGCTGCGGCGCTCTCGACACTGCTATGAGATTC 412				
106	GGCTGCAACTCTCTCGAGAGCATGCTTCAAAAAGACATCAATTCCGACATATCTATACCT 165				
411	GGCTGCAACTCTCTCGAGAGCATGCTTCAAAAAGACATCAATTCCGACATATCTATACCT 352				
166	GAATACAAAGAGCTTCTCAAGAGCTTCAAGAGAGCTGATGCGCGTGCAGAGAGCTATGGG 225				
351	GAATACAAAGAGCTTCTCAAGAGCTTCAAGAGAGCTGATGCGCGTGCAGAGAGCTATGGG 292				
226	AAATTCAAGAGCTTCTCTCAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 285				

Db	291	AAATGACAGAGTGTTCCTCAGCCAGTCACATGAAAGCTGTGAAAAAAGCTTGGACGTG	232
Qy	286	ATGCATACAGTGTACACAGCATTTGGTGTAAATGTGAAGAGTAATTAACCTTACCCAGC	345
Db	231	ATGCATACAGTGTACACAGCATTTGGTGTAAATGTGAAGAGTAATTAACCTTACCCAGC	172
Qy	346	CGTTGGCTCAGAGGCTCAGACATATGGCCAGAACTCATCTGTGATTGCTGGAACCA	405
Db	171	CGTTGGCTCAGAGGCTCAGACATATGGCCAGAACTCATCTGTGATTGCTGGAACCA	112
Qy	406	CTTCTCTCTGTGTC--TTGCTTTTATGTGGAACTGCTGACACACTGTTGAAAGCTCA	462
Db	111	CTTCTCTCTGTGTC--TTTATGTGGAACTGCTGACACACTGTTGAAAGCTCA	52
RESULT 9			
LOCUS	AM779377/c	458 bp	mRNA linear EST 12-MAY-2000
DEFINITION	hm79908.x1 NCI-CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3034142 3		
	similar to TR:075556 075556 MAMMAGLOBIN B PRECURSOR. ; mRNA		
ACCESSION	AM779377		
VERSION	AM779377.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
COMMENT	Tumor Gene Index Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cga@bcrf-remail.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov Seq primer: -40up from gldco High quality sequence stop: 458. Location/Qualifiers 1.458 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3034142" /clone_lib="NCI-CGAP_Kid11" /lab_host="DH10B" /note="Organ: kidney; Vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP_Kid3 was prepared, and ss circles were made in vitro. Following RAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Peltina Bonaldo.		
BASE COUNT	132 a	94 c	99 g 132 t 1 others
ORIGIN			
Query Match	81.7%	Score 389;	DB 10; Length 458;
Best Local Similarity	96.7%	Pred. No. 1.1e-100;	
Matches 408;	Conservative	0; Mismatches 11;	Indels 3; Gaps 1;
Qy	46	ATGACAGTGTATGTGTCTCTATGTGTGGCGCCCTCTCTCGACAGTGTATGACATTCT	105
Db	458	ATGACAGTGTATGTGTCTCTATGTGTGGCGCCCTCTCTCGACAGTGTATGACATTCT	399

OY 106 GGCCTCAACCTCCGAGACATCGTTGAAAAGACCATCAATTCGACATATCTATACCT 165
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 398 GGCCTCAACCTCCGAGACATCGTTGAAAAGACCATCAATTCGACATATCTATACCT 339
 OY 166 GAATCAAAAGAGCTTCTTCAAGATTCAATAGACATGATGCCGTGCAGAGGCTATGGGG 225
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 338 GAATCAAAAGAGCTTCTTCAAGATTCAATAGACATGATGCCGTGCAGAGGCTATGGGG 279
 OY 226 AATTCGAAGCAGTGTTCCTTCACACACTCATAGAACTGTGAAAACTTTGGACTGATG 285
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 278 AATTCGAAGCAGTGTTCCTTCACACACTCATAGAACTGTGAAAACTTTGGACTGATG 219
 OY 286 ATGCATACAGTGTACGACACATTTGGTGAATATGAAAGATTAATTAATTTACCAAGG 345
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 218 ATGCATACAGTGTACGACACATTTGGTGAATATGAAAGATTAATTAATTTACCAAGG 159
 OY 346 CGTTGGCTCAGAGGCTACAGACTATGGCCAGAACTCATCTGTGATTGCTAGAACCA 405
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 158 CGTTGGCTCAGAGGCTACAGACTATGGCCAGAACTCATCTGTGATTGCTAGAACCA 99
 OY 406 CTTTCTTCTTGTG---TTGCTTTTATGTGGAACTGTAGACAACTGTGAAACCTCAA 462
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 98 CTTTCTTCTTGTGTTGTGTCTTTTATGTGGAACACTGTAGACAACTGTGAAACCTCAA 39
 OY 463 TT 464
 Db 38 AT 37
 RESULT 10 458 bp mRNA linear EST 12-MAY-1999
 A1491987
 LOCUS 1007f11.x1 NCI CGAP Ut2 Homo sapiens cDNA clone IMAGE:2178381 3'
 DEFINITION similar to SW:MAMG_HUMAN Q13296 MAMMAGLOBIN PRECURSOR. [1] ;, mRNA
 sequence.
 ACCESSION A1491987
 VERSION A1491987.1 GI:4393001
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 458)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapsb@mail.nih.gov
 Tissue Procurement: Christopher Koskaluk, M.D., Ph.D., Michael R.
 Emmerit-Buck, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bdnp/image/image.html
 Insert length: 577 Std Error: 0.00
 Seq primer: -400p from Glbco
 High quality sequence stop: 399
 POLYA-No.
 FEATURES
 source
 1. 458
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2178381"
 /clone_lib="NCI-CGAP_Ut2"
 /tissue_type="moderately-differentiated endometrial
 adenocarcinoma, 3 pooled tumors"
 /lab_host="DH10B"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1; Salt:
 Site 2: Nott; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.85 kb. Life Technologies catalog #:

BASE COUNT 11539-012*
 ORIGIN
 Query Match 80.9% Score 385.2; DB 9; Length 458;
 Best Local Similarity 96.2%; Pred. No. 1.4e-99;
 Matches 406; Conservative 0; Mismatches 13; Indels 3; Gaps 1;
 OY 46 ATGAAGCTGCTGATGCTGCTTCATGCTGCGCGCCCTCTCTCTGACTGTATGACATTTCT 105
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 458 ATGAAGCTGCTGATGCTGCTTCATGCTGCGCGCCCTCTCTCTGACTGTATGACATTTCT 399
 OY 106 GGCCTCAACCTCCGAGAGACATGTTGAAAAGACCATCAATTCGACATATCTATACCT 165
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 398 GGCCTCAACCTCCGAGAGACATGTTGAAAAGACCATCAATTCGACATATCTATACCT 339
 OY 166 GAATCAAAAGAGCTTCTTCAAGATTCAATAGACATGATGCCGTGCAGAGGCTATGGGG 225
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 338 GAATCAAAAGAGCTTCTTCAAGATTCAATAGACATGATGCCGTGCAGAGGCTATGGGG 279
 OY 226 AATTCGAAGCAGTGTTCCTTCACACACTCATAGAACTGTGAAAACTTTGGACTGATG 285
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 278 AATTCGAAGCAGTGTTCCTTCACACACTCATAGAACTGTGAAAACTTTGGACTGATG 219
 OY 286 ATGCATACAGTGTACGACACATTTGGTGAATATGAAAGATTAATTAATTTACCAAGG 345
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 218 ATGCATACAGTGTACGACACATTTGGTGAATATGAAAGATTAATTAATTTACCAAGG 159
 OY 346 CGTTGGCTCAGAGGCTACAGACTATGGCCAGAACTCATCTGTGATTGCTAGAACCA 405
 ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 Db 158 CGTTGGCTCAGAGGCTACAGACTATGGCCAGAACTCATCTGTGATTGCTAGAACCA 99
 OY 406 CTTTCTTCTTGTG---TTGCTTTTATGTGGAACTGTAGACAACTGTGAAACCTCAA 462
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 98 CTTTCTTCTTGTGTTGTGTCTTTTATGTGGAACACTGTAGACAACTGTGAAACCTCAA 39
 OY 463 TT 464
 Db 38 AT 37
 RESULT 11 396 bp mRNA linear EST 18-APR-1997
 A0297402
 LOCUS EST112936 Endometrial tumor Homo sapiens cDNA 5' end similar to
 DEFINITION similar to steroid-binding protein, C3 chain, prostate, mRNA
 sequence.
 ACCESSION A0297402
 VERSION A0297402.1 GI:1949735
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 396)
 Adams, M.D., Kerlanage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult
 C., J., Lee, N.H., Kirkness, E.F., Weinstein, K.G., Gocke, J.D., White
 O., Sutton, G., Blake, J.A., Brandon, R.C., Man, Wai, C., Clayton, R.A.,
 Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald,
 L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Glodok, A.,
 Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M.,
 Kelley, J.C., Liu, L.-I., Marmaro, S.M., Merick, J.M.,
 Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,
 Phillips, C.A., Ryder, S.E., Scott, J.L., Saudet, D.M., Shirley, R.,
 Small, K.V., Spriggs, T.A., Utecht, T.R., Weidman, J.F., Li, Y.,
 Benharik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,
 Dinke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W.,
 Hu, J.S., Greene, J.M., Hudson, P., Kim, A.K., Kozak, D.L.,
 Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L.,
 Wei, Y.F., Wang, J., Xu, C., Yu, G.-L., Ruben, S.M., Dillon, P.J., Fannon
 M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and
 Venter, J.C.
 Initial assessment of human gene diversity and expression patterns

LOCUS AAA93295 435 bp mRNA linear EST 19-AUG-1997
 DEFINITION ng97e06.s1 NCI_CGAP_Thy1 Homo sapiens cDNA clone IMAGE:942754
 similar to TR:G1199396 G1199396 MAMMAGLOBIN.; mRNA sequence.
 ACCESSION AAA93295
 VERSION AAA93295.1 GI:2223136
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 435)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaps-remail.nih.gov
 Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: David B. Krizman, Ph.D.
 DNA sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/dbp/image/image.html
 Insert length: 887 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 401.
 Location/Qualifiers
 1..435
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="942754"
 /clone_lib="NCI_CGAP_Thy1"
 /tissue_type="thyroid"
 /lab_host="DH10B"
 /note="Vector: PAMP10; mRNA made from invasive thyroid
 tumor, cDNA made by oligo-dT priming. Non-directionally
 cloned. Size selected on agarose gel, average insert size
 600 bp. Reference: Krizman et al. (1996) Cancer Research
 56:5380-5383."
 BASE COUNT 122 a 86 c 94 g 133 t
 ORIGIN
 Query Match 76.6%; Score 364.6; DB 9; Length 435;
 Best Local Similarity 97.0%; Pred. No. 1.1e-93;
 Matches 383; Conservative 0; Mismatches 9; Indels 3; Gaps 1;
 QY 73 GGGGCGCTCCCTCGCACTGCTATGACATTCGCTGCGCAACTCTCTGAGACATGCTT 132
 DB 435 GGGGCGCTCCCTCGCACTGCTATGACATTCGCTGCGCAACTCTCTGAGACATGCTT 376
 QY 133 GAAAGACCATCAATTCGACATATCTATACGTAATACAAAGAGCTTCTCAAGAGTTC 192
 DB 375 GAAAGACCATCAATTCGACATATCTATACGTAATACAAAGAGCTTCTCAAGAGTTC 316
 QY 375 ATAGACATGATGCGCGTGCAGAGCTATGCGGAATTCAGAGTGTTCCTCAACGAG 252
 DB 193 ATAGACATGATGCGCGTGCAGAGCTATGCGGAATTCAGAGTGTTCCTCAACGAG 256
 QY 315 ATAGACATGATGCGCGTGCAGAGCTATGCGGAATTCAGAGTGTTCCTCAACGAG 256
 DB 253 TCACATGAACTGTGAAAACTTTGACATGATGATGATGATGATGATGATGATGATG 312
 QY 255 TCACATGAACTGTGAAAACTTTGACATGATGATGATGATGATGATGATGATGATG 196
 DB 313 TGTATATGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 372
 QY 195 TGTATATGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 136
 QY 373 GCCCAAGACTCATCTGTTGATTGCTAGAAAACCACTTCTTCTTG -TTCCTTTTAT 429
 DB 135 GCCCAAGACTCATCTGTTGATTGCTAGAAAACCACTTCTTCTTGTTGTTCTTTTAT 76
 QY 430 GTGGGAACCTGTAGACAACCTGTGAACCTCAATT 464

Db 75 GTGGGAACCTGTAGACAACCTGTGAACCTCAATT 41
 RESULT 14
 BE044895
 LOCUS
 DEFINITION BE044895 365 bp mRNA linear EST 08-JUN-2000
 h080303.x1 NCI_CGAP_Thy6 Homo sapiens cDNA clone IMAGE:3021509 3'
 similar to TR:075556 075556 MAMMAGLOBIN B PRECURSOR.; mRNA
 sequence.
 ACCESSION BE044895
 VERSION BE044895.1 GI:8361948
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 365)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgaps-remail.nih.gov
 Tissue Procurement: Klaus Kaserer M.D., Bruno Niederle M.D., Mike
 Emmert-Buck M.D., Ph.D., Vlado Knezevic M.D.
 cDNA Library Preparation: Krizman Laboratory
 DNA sequencing by: The I.M.A.G.E. Consortium/LLNL
 Clone distribution: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
 info@image.llnl.gov
 Possible reversed clone: similarity on wrong strand
 Possible reversed clone: polyt not found
 Seq primer: -40UP from Gibco
 High quality sequence stop: 364.
 Location/Qualifiers
 1..365
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="3021509"
 /clone_lib="NCI_CGAP_Thy6"
 /tissue_type="normal epithelium"
 /lab_host="DH10B"
 /note="Organ: thyroid; Vector: PAMP10; mRNA made from
 normal thyroid epithelium, cDNA made by oligo-dT
 priming. Non-directionally cloned into UDG sites. Size-selected on
 agarose gel, average insert size 500 bp. Primary library.
 cDNA Library Preparation: David B. Krizman, Ph.D.
 REFERENCE: Krizman et al. (1996) Cancer Research
 56:5380-5383."
 BASE COUNT 103 a 85 c 85 g 92 t
 ORIGIN
 Query Match 75.0%; Score 357; DB 10; Length 365;
 Best Local Similarity 100.0%; Pred. No. 1.6e-91;
 Matches 357; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 50 AGCTCTATGCTCTCATGCTGCGGCGCTCTCTCTGCACTGCTATGCAATTCGCT 109
 DB 9 AGCTCTATGCTCTCATGCTGCGGCGCTCTCTCTGCACTGCTATGCAATTCGCT 68
 QY 110 GCAAACTCTGAGACATGTTGTAAGAAAGCAATTCATTCGCAATTCATACCTGAT 169
 DB 69 GCAAACTCTGAGACATGTTGTAAGAAAGCAATTCATTCGCAATTCATACCTGAT 128
 QY 170 ACAAGAGCTTCTTCAGAGTTCATAGACAGTATGCGCGTGCAGAGCTATGGGAAT 229
 DB 129 ACAAGAGCTTCTTCAGAGTTCATAGACAGTATGCGCGTGCAGAGCTATGGGAAT 188
 QY 230 TCACAGAGTTCCTCTCAACGACATGAACTGTGAAAACTTTGAGCTGATGATGC 289
 DB 189 TCACAGAGTTCCTCTCAACGACATGAACTGTGAAAACTTTGAGCTGATGATGC 248

```

QY 290 ATACAGTGTACGACAGCATTTGGTGTATATGAAGAGTAATTAACCTTTACCCAGGCGTT 349
      |||||||
DB 249 ATACAGTGTACGACAGCATTTGGTGTATATGAAGAGTAATTAACCTTTACCCAGGCGTT 308
      |||||||
QY 350 TGGCTCAGAGGGCTACAGACTATGCGCAGAACTATCTGTTGATGCTGAAGAACCC 406
      |||||||
DB 309 TGGCTCAGAGGGCTACAGACTATGCGCAGAACTATCTGTTGATGCTGAAGAACCC 365
      |||||||

RESULT 15
BE044893 351 bp mRNA linear EST 08-JUN-2000
LOCUS hnb8c04.x1 NCI-CGAP-Thy6 Homo sapiens cDNA clone IMAGE:3021510 3'
DEFINITION similar to TR:075556 075556 MAMMALOBLIN B PRECURSOR. ; mRNA
sequence.
ACCESSION BE044893
VERSION BE044893
KEYWORDS EST. GI:8361946
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 351)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Klaus Kaserer M.D., Bruno Niederle M.D., Mike
Emmert-Buck M.D., Ph.D., Vlado Knezevic M.D.
cDNA Library Preparation: Krizman Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Possible reversed clone: similarity on wrong strand
Possible reversed clone: polyT not found
Seq primer: -40UP from Gibco
High quality sequence stop: 351.

FEATURES
Source
1..351
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3021510"
/clone_id="NCI-CGAP Thy6"
/tissue_type="normal epithellum"
/lab_host="DH10B"
/note="Organ: thyroid; Vector: pAMP10; mRNA made from
normal thyroid epithellum; cDNA made by oligo-dT priming.
Non-directionally cloned into UDG sites. Size-selected on
agarose gel, average insert size 500 bp. Primary library.
cDNA Library Preparation: David B. Krizman, Ph.D.
REFERENCE: Krizman et al. (1996) Cancer Research
56:5380-5383."

BASE COUNT 98 a 81 c 83 g 89 t
ORIGIN

Query Match 72.1%; Score 343; DB 10; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.6e-87;
Matches 343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 ACCTGCTATGCTCTCTGCTGCGGCCCTCCCTGCGACATGCTATGCGAGATTCGGCT 109
      |||||||
DB 9 ACCTGCTATGCTCTCTGCTGCGGCCCTCCCTGCGACATGCTATGCGAGATTCGGCT 68
      |||||||
QY 110 GCAAACTCCTGAGAGACATGGTTGAAAGACCATTCATTCGACATATCTATACCTGAAT 169
      |||||||
DB 69 GCNAACTCCTGAGAGACATGGTTGAAAGACCATTCATTCGACATATCTATACCTGAAT 128
      |||||||
QY 170 ACAAGAGCTTCTTCAAGAGTTCAATAGACGTATGCGCGCTGACAGGCTATGGGAAT 229

```

```

DB 129 ACAAGAGCTTCTTCAAGAGTTCAATAGACGTATGCGCGCTGACAGGCTATGGGAAT 188
      |||||||
QY 230 TCACAGAGTGTTCCTCAACCAAGTCACATGAACCTGAAAACTTTGGAGCTGATGATGC 289
      |||||||
DB 189 TCACAGAGTGTTCCTCAACCAAGTCACATGAACCTGAAAACTTTGGAGCTGATGATGC 248
      |||||||
QY 290 ATACAGTGTACGACAGCATTTGGTGTATATGAAGAGTAATTAACCTTTACCCAGGCGTT 349
      |||||||
DB 249 ATACAGTGTACGACAGCATTTGGTGTATATGAAGAGTAATTAACCTTTACCCAGGCGTT 308
      |||||||
QY 350 TGGCTCAGAGGGCTACAGACTATGCGCAGAACTATCTGTTGA 392
      |||||||
DB 309 TGGCTCAGAGGGCTACAGACTATGCGCAGAACTATCTGTTGA 351
      |||||||

Search completed: January 25, 2003, 13:18:08
Job time : 2260 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 25, 2003, 10:29:19 : Search time 3276 Seconds
(without alignments)
4228.613 Million cell updates/sec

Title: US-09-806-302a-1

Perfect score: 476
Sequence: 1 agcagtcgcacgcacgact.....cctcaatcattcattca 476

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_da: *
2: gb_hcg: *
3: gb_in: *
4: gb_ov: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_lo: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_hcg_hum: *
31: em_hcg_inv: *
32: em_hcg_other: *
33: em_hcg_mus: *
34: em_hcg_pln: *
35: em_hcg_rod: *
36: em_hcg_man: *
37: em_hcg_vir: *
38: em_sy: *
39: em_hcgo_hum: *
40: em_hcgo_mus: *
41: em_hcgo_other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	476	100.0	476	AR142399	AR142399 Sequence
2	476	100.0	476	AR182545	AR182545 Sequence
3	434.2	91.2	496	HS2424173	AJ224173 Homo sapi
4	431.6	90.7	517	AX013093	AX013093 Sequence
5	431.6	90.7	517	AF071219	AF071219 Homo sapi
6	407.4	85.6	491	AX335698	AX335698 Sequence
7	223.8	47.0	495	AR172283	AR172283 Sequence
8	223.8	47.0	503	AR080285	AR080285 Sequence
9	223.8	47.0	503	AR095403	AR095403 Sequence
10	223.8	47.0	503	AX213242	AX213242 Sequence
11	223.8	47.0	503	165735	165735 Sequence 1
12	223.8	47.0	503	HS031447	U33147 Human mamma
13	223.8	47.0	535	AR207542	AR207542 Sequence
14	222.2	46.7	503	AX283024	AX283024 Sequence
15	211.6	44.1	429	AX213264	AX213264 Sequence
16	210	44.1	429	AX213261	AX213261 Sequence
17	208.4	43.8	429	AX213258	AX213258 Sequence
18	208.4	43.8	429	AX213259	AX213259 Sequence
19	208.4	43.8	429	AX213260	AX213260 Sequence
20	208.4	43.8	429	AX213263	AX213263 Sequence
21	205.6	43.2	388	AX213235	AX213235 Sequence
22	202.4	42.5	388	AX213234	AX213234 Sequence
23	202.4	42.5	388	AX213236	AX213236 Sequence
24	202.4	42.5	388	AX213237	AX213237 Sequence
25	202.4	42.5	388	AX213238	AX213238 Sequence
26	202.4	42.5	388	AX213239	AX213239 Sequence
27	202.4	42.5	388	AX213240	AX213240 Sequence
28	202.4	42.5	388	AX213241	AX213241 Sequence
29	201.2	42.3	403	AR080288	AR080288 Sequence
30	201.2	42.3	403	AR095406	AR095406 Sequence
31	201.2	42.3	403	165738	165738 Sequence 5
32	195.2	41.0	467	AX156186	AX156186 Sequence
33	191.8	40.3	172281	AC074200	AC074200 Homo sapi
34	191.8	40.3	173348	AF003306	AF003306 Homo sapi
35	191.8	40.3	178770	AC090725	AC090725 Homo sapi
36	191.8	40.3	181140	AP002793	AP002793 Homo sapi
37	188.4	39.6	420	AX213265	AX213265 Sequence
38	188.4	39.6	420	AX213270	AX213270 Sequence
39	186.8	39.2	420	AX213262	AX213262 Sequence
40	182.4	38.3	379	AX213233	AX213233 Sequence
41	172	36.1	506	AF308618	AF308618 Oryctolag
42	169.2	35.5	461	AF308620	AF308620 Oryctolag
43	169	35.5	486	AX358804	AX358804 Sequence
44	169	35.5	486	AX362297	AX362297 Sequence
45	160.8	33.8	452	AF308619	AF308619 Oryctolag

ALIGNMENTS

RESULT 1
LOCUS AR142399
DEFINITION Sequence 5 from patent US 6174992.
ACCESSION AR142399
VERSION AR142399.1 GI:15102699
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 476)
AUTHORS Ni,J., Yu,G.-L. and Gentz/R.
TITLE Human endometrial specific steroid-binding factor I, II and III
JOURNAL Patent: US 6174992-A 5 16-JAN-2001;
FEATURES Location/Qualifiers

source 1. .476 /organism="unknown"
BASE COUNT 131 a 115 c 108 g 125 t
ORIGIN

Query Match 100.0%; Score 476; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 8.7e-131;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAGCTGGCAGGAGCAGTAAACAGACAGAGAGCGGCGTCCGCTGAAGTGTGATG 60
DB 1 AGGAGCTGGCAGGAGCAGTAAACAGACAGAGAGCGGCGTCCGCTGAAGTGTGATG 60
QY 61 GTTCATATGCTGGCGCCCTCTCTCTGCTGCTGATGAGATTCGCTGCAAGCTCTG 120
DB 61 GTTCATATGCTGGCGCCCTCTCTCTGCTGCTGATGAGATTCGCTGCAAGCTCTG 120
QY 121 GAGGACATGTTGAAAAACCATCAATTCGACATATCTATCTACTGTAATACAAAGCTT 180
DB 121 GAGGACATGTTGAAAAACCATCAATTCGACATATCTATCTACTGTAATACAAAGCTT 180
QY 181 GTTCAAGAGTTTCATAGACAGATGATCCGCTGACAGAGCTATGGGAAATTCAGCAGTGT 240
DB 181 GTTCAAGAGTTTCATAGACAGATGATCCGCTGACAGAGCTATGGGAAATTCAGCAGTGT 240
QY 241 TTCTCAACAGAGTCAATAGAGAGTCACTGTAAGAACTTTGGATGATGATGATGATGATG 300
DB 241 TTCTCAACAGAGTCAATAGAGAGTCACTGTAAGAACTTTGGATGATGATGATGATGATG 300
QY 301 GACAGCATTTGGTGTATATATAGAGATTAATTAACCTTCCAGAGCGTTGGCTGACAGAG 360
DB 301 GACAGCATTTGGTGTATATATAGAGATTAATTAACCTTCCAGAGCGTTGGCTGACAGAG 360
QY 361 GGTACAGACTATGGCAGACACTGATCTGTTGATGATGATGATGATGATGATGATGATG 420
DB 361 GGTACAGACTATGGCAGACACTGATCTGTTGATGATGATGATGATGATGATGATGATG 420
QY 421 GCTTTTATGTTGGAGCACTGTACAGAACTGTTGAAACCTCAATTCATTCATTCATTC 476
DB 421 GCTTTTATGTTGGAGCACTGTACAGAACTGTTGAAACCTCAATTCATTCATTCATTC 476

RESULT 2
AR182545 476 bp DNA linear PAT 20-APR-2002
LOCUS AR182545
DEFINITION Sequence 5 from patent US 6338948.
ACCESSION AR182545
VERSION AR182545.1 GI:20225752
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS N.J., Yu, G., L. and Gentz, R.
TITLE Human endometrial specific steroid-binding factor I, II and III
JOURNAL Patent: US 6338948-A 5 15-JAN-2002;
FEATURES
source location/Qualifiers
1. .476
/organism="unknown"

BASE COUNT 131 a 115 c 105 g 125 t
ORIGIN

Query Match 100.0%; Score 476; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 8.7e-131;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGGAGCTGGCAGGAGCAGTAAACAGACAGAGAGCGGCGTCCGCTGAAGTGTGATG 60
DB 1 AGGAGCTGGCAGGAGCAGTAAACAGACAGAGAGCGGCGTCCGCTGAAGTGTGATG 60
QY 61 GTTCATATGCTGGCGCCCTCTCTCTGCTGCTGATGAGATTCGCTGCAAGCTCTG 120
DB 61 GTTCATATGCTGGCGCCCTCTCTCTGCTGCTGATGAGATTCGCTGCAAGCTCTG 120

QY 121 GAGGACATGTTGAAAAACCATCAATTCGACATATCTATCTACTGTAATACAAAGCTT 180
DB 121 GAGGACATGTTGAAAAACCATCAATTCGACATATCTATCTACTGTAATACAAAGCTT 180
QY 181 GTTCAAGAGTTTCATAGACAGATGATCCGCTGACAGAGCTATGGGAAATTCAGCAGTGT 240
DB 181 GTTCAAGAGTTTCATAGACAGATGATCCGCTGACAGAGCTATGGGAAATTCAGCAGTGT 240

QY 241 TTCTCAACAGAGTCAATAGAGAGTCACTGTAAGAACTTTGGATGATGATGATGATGATG 300
DB 241 TTCTCAACAGAGTCAATAGAGAGTCACTGTAAGAACTTTGGATGATGATGATGATGATG 300
QY 301 GACAGCATTTGGTGTATATATAGAGATTAATTAACCTTCCAGAGCGTTGGCTGACAGAG 360
DB 301 GACAGCATTTGGTGTATATATAGAGATTAATTAACCTTCCAGAGCGTTGGCTGACAGAG 360
QY 361 GGTACAGACTATGGCAGACACTGATCTGTTGATGATGATGATGATGATGATGATGATG 420
DB 361 GGTACAGACTATGGCAGACACTGATCTGTTGATGATGATGATGATGATGATGATGATG 420
QY 421 GCTTTTATGTTGGAGCACTGTACAGAACTGTTGAAACCTCAATTCATTCATTCATTC 476
DB 421 GCTTTTATGTTGGAGCACTGTACAGAACTGTTGAAACCTCAATTCATTCATTCATTC 476

RESULT 3
HSR224173 496 bp mRNA linear PRI 25-MAR-1999
LOCUS HSR224173
DEFINITION Homo sapiens mRNA for lipophilin C.
ACCESSION AJ224173
VERSION AJ224173.1 GI:4107232
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Zhao, C., Nguyen, T., Yusifov, T., Glasgow, B.J., and Lehrer, R.I.
TITLE Lipophilins: human peptides homologous to rat prostatein
JOURNAL Biochem. Biophys. Res. Commun. 256 (1), 147-155 (1999)
MEDLINE 99167354
PUBMED 10065439
REFERENCE
AUTHORS Zhao, C.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-1998) Zhao C., UCLA Dept. of Medicine, CHS 37055, Los Angeles, CA 90095, USA
FEATURES
source location/Qualifiers
1. .496
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS
41..328
/codon_start=1
/product="lipophilin C"
/protein_id="CAA11865.1"
/db_xref="GI:4107232"
/translation="MKILVIMIAALILHYADSGCKLLEDMVERKTIINSDISTPEYKE
LLQEFISDAAAEAMKRFKOCFLNDSHRLTKNFGILMHTIYDSTINCKMKNK"
mat_peptide
95..325
/product="lipophilin C"

BASE COUNT 141 a 119 c 102 g 134 t
ORIGIN

Query Match 91.2%; Score 434.2; DB 9; Length 496;
Best Local Similarity 97.8%; Pred. No. 2.6e-118;
Matches 451; Conservative 0; Mismatches 8; Indels 2; Gaps 1;

QY 6 CTGCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 65
DB 1 CTGCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 60
QY 66 CATGCTGGGGGGGCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 125

```

|||||
Db 61 CATGCTGGGCGCCCTCTCTCTGACGCTATGCAATTCCTGGTGAACCTCTGAGAGA 120
OY 126 CATGCTGAAAAAGACCAATTCGCCATATCTATACCTGGAATACAAAGACTTCTTCA 185
Db 121 CATGCTGAAAAAGACCAATTCGCCATATCTATACCTGGAATACAAAGACTTCTTCA 180
OY 186 AGAGTTCAATAGACAGTATGCGCGCTGAGAGGCTATGGGAAATTCAGCACTTTTCT 245
Db 181 AGAGTTCAATAGACAGTATGCGCGCTGAGAGGCTATGGGAAATTCAGCACTTTTCT 240
OY 246 CAACAGTCACATAGAACTCTGAAAACTTTGGACTGATGATCATACAGTACAGACAG 305
Db 241 CAACAGTCACATAGAACTCTGAAAACTTTGGACTGATGATCATACAGTACAGACAG 300
OY 306 CATTTGGTATATATGAAGAGTATTAACCTTACCCAAAGCGCTGGCTCAGAGGGCTAC 365
Db 301 CATTTGGTATATATGAAGAGTATTAACCTTACCCAAAGCGCTGGCTCAGAGGGCTAC 360
OY 366 AGACTATGGCCAGAACTCATCTGTTGATTCCTAGAAACCACTTT-CTTCTTGTGCT 423
Db 361 AGACTATGGCCAGAACTCATCTGTTGATTCCTAGAAACCACTTTCTTCTTGTGCT 420
OY 424 TTTTATGTGGAACTGTAGCACTGTGAAACCTCAAT 464
Db 421 TTTTATGTGGAACTGTAGCACTGTGAAACCTCAAT 461

RESULT 4
AX013093 517 bp DNA linear PAT 07-SEP-2000
LOCUS Sequence 33 from Patent WO954461.
DEFINITION AX013093
ACCESSION AX013093
VERSION AX013093.1 GI:10040259
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 517)
AUTHORS Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pillarsky,C.
TITLE Human nucleic acid sequences of endometrium tumour tissue
JOURNAL Patient: WO 954461-A 33 28-OCT-1999;
SCHMITT AMIN (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METZGEN GES FUER GENOMFORSCHUN
(DE); PILLARSKY CHRISTIAN (DE)
FEATURES
source
1. 517
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 156 a 121 c 106 g 134 t
ORIGIN
Query Match 90.7%; Score 431.6; DB 6; Length 517;
Best Local Similarity 97.4%; Pred. No. 1.5e-117;
Matches 450; Conservative 0; Mismatches 9; Indels 3; Gaps 1;
OY 6 CTGCCACGACGACATGAAAGACAGACAGCCGCTCGCATATACCTGCTATGCTCT 65
Db 1 CTGCCACGACGACATGAAAGACAGACAGCCGCTCGCATATACCTGCTATGCTCT 60
OY 66 CATGCTGGGCGCCCTCTCTCTGCACTGCTATGAGATTCTGGCTGCAAACTCCTGAGGA 125
Db 61 CATGCTGGGCGCCCTCTCTCTGCACTGCTATGAGATTCTGGCTGCAAACTCCTGAGGA 120
OY 126 CATGCTGAAAAAGACCAATTCGCCATATCTATACCTGGAATACAAAGACTTCTTCA 185
Db 121 CATGCTGAAAAAGACCAATTCGCCATATCTATACCTGGAATACAAAGACTTCTTCA 180
OY 186 AGAGTTCAATAGACAGTATGCGCGCTGAGAGGCTATGGGAAATTCAGCACTTTTCT 245
Db 181 AGAGTTCAATAGACAGTATGCGCGCTGAGAGGCTATGGGAAATTCAGCACTTTTCT 240

```

```

OY 246 CAACAGTCACATAGAACTCTGAAAACTTTGGACTGATGATCATACAGTACAGACAG 305
Db 241 CAACAGTCACATAGAACTCTGAAAACTTTGGACTGATGATCATACAGTACAGACAG 300
OY 306 CATTTGGTATATATGAAGAGTATTAACCTTACCCAAAGCGCTGGCTCAGAGGGCTAC 365
Db 301 CATTTGGTATATATGAAGAGTATTAACCTTACCCAAAGCGCTGGCTCAGAGGGCTAC 360
OY 366 AGACTATGGCCAGAACTCATCTGTTGATTCCTAGAAACCACTTTCTTCTGTC--TTGC 422
Db 361 AGACTATGGCCAGAACTCATCTGTTGATTCCTAGAAACCACTTTCTTCTGTCGTC 420
OY 423 TTTTATGTGGAACTGTAGCACTGTGAAACCTCAAT 464
Db 421 TTTTATGTGGAACTGTAGCACTGTGAAACCTCAAT 462

RESULT 5
AF071219 517 bp mRNA linear PRI 26-NOV-1998
LOCUS AF071219
DEFINITION Homo sapiens mamaglobin B precursor, mRNA, complete cds.
ACCESSION AF071219
VERSION AF071219.1 GI:3288867
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 517)
AUTHORS Becker,R.M., Darrow,C., Zimonjic,D.B., Popescu,N.C., Watson,M.A.
and Fleming,T.P.
TITLE Identification of mamaglobin B, a novel member of the uteroglobin
gene family
JOURNAL Genomics 54 (1), 70-78 (1998)
MEDLINE 99026127
PUBMED 9806831
REFERENCE 2 (bases 1 to 517)
AUTHORS Becker,R.M., Darrow,C.M., Watson,M.A. and Fleming,T.P.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1998) Ophthalmology, Washington University School
of Medicine, 660 S. Euclid, St. Louis, MO 63110, USA
FEATURES
source
1. 517
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
65..352
/note="member of the uteroglobin family"
/codon_start=1
/product="mamaglobin B precursor"
/protein_id="AAC79996.1"
/db_xref="GI:3288868"
/translation="MKLLVIMLAALILHCYADSGCKLLEDVVEKTIINSISPEYKE
LLDFISDAAAEAMKRFKOCFLFNOSHRLLKFLGLMHTVDSIWMKNSN"
BASE COUNT 144 a 129 c 105 g 139 t
ORIGIN
Query Match 90.7%; Score 431.6; DB 9; Length 517;
Best Local Similarity 97.4%; Pred. No. 1.5e-117;
Matches 450; Conservative 0; Mismatches 9; Indels 3; Gaps 1;
OY 6 CTGCCACGACGACATGAAAGACAGACAGCCGCTCGCATATACCTGCTATGCTCT 65
Db 25 CTGCCACGACGACATGAAAGACAGACAGCCGCTCGCATATACCTGCTATGCTCT 84
OY 66 CATGCTGGGCGCCCTCTCTCTGCACTGCTATGAGATTCTGGCTGCAAACTCCTGAGGA 125
Db 85 CATGCTGGGCGCCCTCTCTCTGCACTGCTATGAGATTCTGGCTGCAAACTCCTGAGGA 144
OY 126 CATGCTGAAAAAGACCAATTCGCCATATCTATACCTGGAATACAAAGACTTCTTCA 185

```


LOCUS	AR080285		503 bp.	DNA	linear	PAT 31-AUG-2000
DEFINITION	Sequence 1 from patent US 5968754.					
ACCESSION	AR080285					
VERSION	AR080285.1	GI:10007020				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	Unclassified.					
AUTHORS	Watson,M.A. and Fleming,T.P.					
TITLE	Mammaglobin, a mammary-specific breast cancer protein					
JOURNAL	Patent: US 5968754-A 1 19-OCT-1999;					
FEATURES	Location/Qualifiers					
source	1..503					
	/organism="unknown"					
BASE COUNT	146 a	118 c	97 g	142 t		
ORIGIN						
Query Match	47.0%;	Score 223.8;	DB 6;	Length 503;		
Best Local Similarity	71.4%;	Pred. No. 1.4e-55;				
Matches 325; Conservative	0;	Mismatches 122;	Indels	8;	Gaps	2;
Oy	7 TGCACGACGACTGTAAGCACAACAGACAGCCGCCCGCATGAAGTGTGTGATGTC	66				
Dd	22 TGGCACCGCGGATGAAACACGACGACGACBCCACACATGAAGTGCTGATGTC	81				
Oy	67 ATGCTGGCGGCTCTCTCTGCACTGCTATGCAATTCTGCTGCACAACTCTGCAGAC	126				
Dd	82 ATGCTGGCGGCTCTCTCTGCACTGCTATGCAATTCTGCTGCACAACTCTGCAGAC	141				
Oy	127 ATGCTGAAAAGACATCAATTCGACATATCTATACCTGAATCAAGAAGCTTTCCA	186				
Dd	142 GTGATTTCCAGACATCAATCCACAGTGTCTAAGACTGATATCAAGAAGCTTTCCA	201				
Oy	187 GAGTCAATAACAGTAGTATGCGCTGCAGAGCTATGGGAAATTCAGCAGTGTTC	246				
Dd	202 GAGTCAATAACAGCATGCTCCACTCAAAATGCCATGATGAATTGAAGGAATGTTCT	261				
Oy	247 AACCAATCAGATGAACCTGAAAAAATTTGGACTGATGATGATGATGATGATGATG	306				
Dd	262 AACCAACGATGAACCTGAGCAATGTTGGATGATGATGATGATGATGATGATGATG	321				
Oy	307 ATTGGCTGAATGTAAGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAAT	366				
Dd	322 AGCTTTGTGATTT-----ATTTAATTAATTAATTAATTAATTAATTAATTAAT	375				
Oy	367 GACTATGGCCGAGACTCATCTGTGATGCTGAGAAC--CACTTCTCTTGTGTGCTT	424				
Dd	376 GGATATGATGAGAAACCACTACGGATGTGCTCAAAACACACCTTCTTTTATGTCT	435				
Oy	425 TTATATGGGAACTGCTAGCAACAATGTTGAACCT	459				
Dd	436 TTTTACTACAACTACAAACAACAATGTTGAACCT	470				
RESULT 9						
LOCUS	AR095403		503 bp.	DNA	linear	PAT 08-SEP-2000
DEFINITION	Sequence 1 from patent US 6004756.					
ACCESSION	AR095403					
VERSION	AR095403.1	GI:10023248				
KEYWORDS						
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	Unclassified.					
AUTHORS	1 (bases 1 to 503)					
TITLE	Method for detecting the presence of breast cancer by detecting an					
JOURNAL	increase in mammaglobin mRNA expression					
FEATURES	Patent: US 6004756-A 1 21-DEC-1999;					
source	Location/Qualifiers					
	1..503					

BASE COUNT	146 a	118 c	97 g	142 t	ORIGIN
Query Match	47.0%	Score 223.8;	DB 6;	Length 503;	
Best Local Similarity	71.4%;	Pred. No. 1.4e-55;			
Matches 325;	Conservative	0;	Mismatches 122;	Indels 8;	Gaps 2;
QY	7	TGCGACCGCAGCATGACAGACAGACGCGCCGCTCCCATGTAAGCTGCTGATGTCCTC	66		
DB	22	TGCGACCGCGAGACTGAACACGACAGCAGCGCCCTCACCATGAAGTTGCTGATGTCCTC	81		
QY	67	ATGCTGGCGCCCTCCTCCTCGACATGCTATGAGATTCCTGGCTGCAAACTCTGAGAGAC	126		
DB	82	ATGCTGGCGCCCTCCTCCTCGACATGCTATGAGATTCCTGGCTGCAAACTCTGAGAGAT	141		
QY	127	ATGCTGAAAGACCATCAATTCGCACATATCTATACCTGATATCAAAAGCTCTTCGA	186		
DB	142	GTGATTTCCACAGCAATCATCCACAAGTGTCTAAGACTGAATCAAAAGAACTTCTCA	201		
QY	187	GAGTTCTATGACAGTATGATCCGCTCGACAGAGCTATGAGGAAATTCAGCAGTGTCTCTC	246		
DB	202	GAGTTCTATGACAGCAATGATCCACTCAAAATGCAATGATGAATGAAGAAATTTTCTT	261		
QY	247	AACCATGACATAGACTCTGAAAACTTTGGACTGATGATGATACAGTATGACGACAC	306		
DB	262	AACCAAGGATTAACCTCTGACGACATGTTAGGGTTATGCAATTAATATATGACAC	321		
QY	307	ATTTGTGTATATGAAAGATTAATTAATTAACCAAGGCGCTTGGCTCAGAGGCTTACA	366		
DB	322	AGCTTTGTGATTT-----ATTTAACTTCTCTGCAAGACCTTGGCTCAGAGACTGCA	375		
QY	367	GAGATGGCCAGAACTCATCTGTATATGCTATGAAG--CACCTTCTCTTGTGTGCTT	424		
DB	376	GCGATGCTAGAGAAACCACTACGATTTGCTTCACAAACCAACACTTCTTCTTATGTCT	435		
QY	425	TTTATGTGGAGACTGCTAGACAACATGTTGAACCT	459		
DB	436	TTTACTACAACTACACAGCAATTTGGAACCT	470		
RESULT 10					
AX213242					
LOCUS	AX213242	503 bp	DNA	linear	PAT 06-SEP-2001
DEFINITION	Sequence 27 from Patent WO0158947.				
ACCESSION	AX213242				
VERSION	AX213242.1	GI:15524172			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eskaya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 503)				
AUTHORS	Carlier, D.; Vedyack, T.S.; Vallieve-Douglas, J.; Houghton, R.L. and Dillon, D.C.				
TITLE	Lipophilin complexes for use in cancer diagnosis and therapy				
JOURNAL	Patent: WO 0158947-A 27 16-AUG-2001;				
FEATURES	CORONA CORPORATION (US)				
SOURCE	Location/Qualifiers				
	1..503				
BASE COUNT	146 a	118 c	97 g	142 t	
ORIGIN	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
Query Match	47.0%	Score 223.8;	DB 6;	Length 503;	
Best Local Similarity	71.4%;	Pred. No. 1.4e-55;			
Matches 325;	Conservative	0;	Mismatches 122;	Indels 8;	Gaps 2;
QY	7	TGCGACCGCAGCATGACAGACAGACGCGCCGCTCCCATGTAAGCTGCTGATGTCCTC	66		
DB	22	TGCGACCGCGAGACTGAACACGACAGCAGCGCCCTCACCATGAAGTTGCTGATGTCCTC	81		
QY	67	ATGCTGGCGCCCTCCTCCTCGACATGCTATGAGATTCCTGGCTGCAAACTCTGAGAGAC	126		
DB	82	ATGCTGGCGCCCTCCTCCTCGACATGCTATGAGATTCCTGGCTGCAAACTCTGAGAGAT	141		
QY	127	ATGCTGAAAGACCATCAATTCGCACATATCTATACCTGATATCAAAAGCTCTTCGA	186		
DB	142	GTGATTTCCACAGCAATCATCCACAAGTGTCTAAGACTGAATCAAAAGAACTTCTCA	201		
QY	187	GAGTTCTATGACAGTATGATCCGCTCGACAGAGCTATGAGGAAATTCAGCAGTGTCTCTC	246		
DB	202	GAGTTCTATGACAGCAATGATCCACTCAAAATGCAATGATGAATGAAGAAATTTTCTT	261		
QY	247	AACCATGACATAGACTCTGAAAACTTTGGACTGATGATGATACAGTATGACGACAC	306		
DB	262	AACCAAGGATTAACCTCTGACGACATGTTAGGGTTATGCAATTAATATATGACAC	321		
QY	307	ATTTGTGTATATGAAAGATTAATTAATTAACCAAGGCGCTTGGCTCAGAGGCTTACA	366		
DB	322	AGCTTTGTGATTT-----ATTTAACTTCTCTGCAAGACCTTGGCTCAGAGACTGCA	375		
QY	367	GAGATGGCCAGAACTCATCTGTATATGCTATGAAG--CACCTTCTCTTGTGTGCTT	424		
DB	376	GCGATGCTAGAGAAACCACTACGATTTGCTTCACAAACCAACACTTCTTCTTATGTCT	435		
QY	425	TTTATGTGGAGACTGCTAGACAACATGTTGAACCT	459		
DB	436	TTTACTACAACTACACAGCAATTTGGAACCT	470		

Qy	67	ATCCTGGGGCCCTCCTCCTGCAGTGCATGCATCAATTCCTGGCTGCACATCTCGAGAGAC	126
Db	82	ATGTGGGGCCCTCTCCAGCAGCTGACGAGAGGCTCTGGCTCTATTTGAGAGAT	141
Qy	127	ATGTTGAAAAGACATCAATTCGACATATCTATACCTGANTACAAAGCTTCTTCA	186
Db	142	GTAATTTTCCAGCAATCAATCACACAGCTGTCTAAAGATGATATCAAAAGAACTTCTTCA	201
Qy	187	GAGTTCAATACAGAGTATGCCGCTCGAGAGGCTATGGGAAATTCAGCAGTCTTCTC	246
Db	202	GAGTTCTAATACACACANTCCCACTCAAAATGCCATAGATGAATTCAGAAATCTTCTT	261
Qy	247	AACAGTCATCATGAACCTGSAAAACTTTGGACGTGATGCAATACAGTGTGCGACAC	306
Db	262	AACCAAAAGGATTAACACTGTGAGCAATGTGGAGTGTATTGCAATTAATATGCAAC	321
Qy	307	ATTTGGTGAATATGAGAGATTAATTAACCTTAACCCAAAGGCTTTGGCTCAGAGGCTTACA	366
Db	322	AGCTTTTGATTTT-----ATTTTAACCTTCTCGAGAGACCTTTGGCTCACAAACTGTGA	375
Qy	367	GACTATGGCAGAACTCATCTGTTAATTTGTCAGAAAC--CACTCTCTTCTGTGTGTT	424
Db	376	GGGATGTGTGAGAAACCACTACGGAATTTGTGTGCAAAACCAACACTTCTCTTATTGTCT	435
Qy	425	TTTATGTGGGAACCTGCTACACAACTGTGAAACCT	459
Db	436	TTTTTACTACAAACTACAGACAATTTGGAACCT	470

RESULT	11				
LOCUS	165735				
DEFINITION	Sequence 1 from patent US 5668267.	503 bp.	DNA	linear	PAT 07-OCT-1997
ACCESSION	165735				
VERSION	165735.1				
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 503)				
TITLE	Watson,M.A. and Fleming,T.P. Polynucleotides encoding mammagloblin, a mammary-specific breast cancer protein				
JOURNAL	Patent: US 5668267-A 1 16-SEP-1997;				
FEATURES	Location/Qualifiers				
Source	1..503				
	/organism="unknown"				
BASE COUNT	146 a 118 c 97 g 142 t				
ORIGIN					

Query Match	47.08;	Score 223.8;	DB 6;	Length 503;
Best Local Similarity	71.48;	Pred. No. 1.4e-55;		
Matches 325;	Conservative	0;	Mismatches 122;	Indels 8;
				Gaps 2;

QY	7	TGGCAGCGACGAGTATACACACGAGAGAGCGCGCTGGCCATCAAGCTCTGATGGTCTC	66
Db	22	TGGCACCCCCGAGCTAACAACCGACGAGGAGCGCTCCACATGAAGTTGCTGATGGTCTC	81
QY	67	ATGCTGGCGGCGCTCTCTCTGCTGACTGCTATGSCAGATTCTGGCTGCGAAACCTCTGAGGAC	126
Db	82	ATGCTGGCGGCGCTCTCTCTGCTGACTGCTATGSCAGATTCTGGCTGCGAAACCTCTGAGGAC	141
QY	127	ATGSGTGGAAAGACCATCAATATCCGACATATCATATACGCGAATACAAAGAGACTCTTCAA	186
Db	142	GTCATTTCCGAAGCAATCATCATTCGCAAGAGTGTCTAGACGATATACAAAGACTCTTCAA	201
QY	187	GAACTTCATGACAGATGATATGCCGCTGTCAGAGGCTATGGGGAATTTCAAGCATGTTTCTC	246
Db	202	GAGTTCATTTAGACAGCAATAGGCCATCTACAAATAGGCCATAGATGAATTAAGAGAAATGTTTCTT	261
QY	247	AACCACTCAATAGAACTGTGAAAAACCTTTGGACTGATGATGCATACAGTGTACGACAGC	306
Db	262	AACCAAAACGATGAACATCTTAGGCAATCTTTGAGGTGTTTATCAATTAATTAATGACAGC	321

Qy	307	ATTTGGTGTAAATTAACAAGTAATTAACCTTTACCACAAGGCGTTTGCGTCAAGGGCTTACA	366
Db	322	AGCCTTGTGATTT-----ATTTAACTTTTCGCAAGACCTTGGCGTCACAGAACTGCA	375
Qy	367	GACTATGGGCACAACACTGCATTCGTGATTTGGTATACAAAC--CACTTCTCTCTGTGTGCTT	424
Db	376	GGGTATGCGTACAAACCAACTACGAGATTCCTGCAAACCAACACTTCTCTTCTATATGCT	435
Qy	425	TTTATGTGGGAACCTGCTACACAACCTGTTGAAACCT	459
Db	436	TTTTACTACAAACTACAAAGACAAATTTGTGGAACCT	470

RESULT 12					
Locus	HSU033147				
Definition	Human mammaglobin mRNA.	503 bp	-mRNA		
Accession	U33147		complete cds.		
Version	U33147.1	GI:1199595		linear	PRI 22-FEB-1996
Keywords					
Source	Homo sapiens.				
Organism	Homo sapiens				

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 503)
Watson, M.A. and Fleming, T.P.
Mammaglobin, a mammary-specific member of the uteroglobin gene
family, is overexpressed in human breast cancer
Cancer, Res. 56 (4), 860-865 (1996)
9622698
8631025
2 (bases 1 to 503)
Watson, M.A.
Direct Submission
Submitted (03-AUG-1995) Mark A. Watson, Washington Univ., School of
Medicine, Dept. Ophthalmology and Visual Sciences, Dept. Genetics,
Div. Laboratory Medicine, Box 8118, 660 S. Euclid Avenue, St.
Louis, MO 63110, USA

FEATURES	Location/Qualifiers
source	1. .503

CDS

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/rtissue_type="breast"
61..342
/note="similar to uteroglobin: Swissprot Accession Number
P02779 and rat prostatic steroid binding protein subunit
C3: Swissprot Accession Number P02780"
/codon_start=1
/product="mamaglobin"
/protein_id="AAC50375.1"
/db_xref="GI:1199596"
/translation="MKLLVLAALSOHVAGSGGLENIYISKINPOVSKREYNE
LLEQIEDNATNNAIDELKRCFLNIDETILSNVEFMQLTYSICLDF"

```

Query Match	Score	DB	Length
47.08;	223.8;	9;	503;

MacCormack	322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000	Conservative	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154,
------------	--	--------------	--

[illegible]

OY	187	GACTTATATGACAGAGTATGCTCCCTGCAGAGCTATGGGAAATTTCAAGCGTCTTCC	246
Db	202	GACTTATATGACAGACATGCCACTCAAAATGCCATAGATGAATGAAGGAATGTTTCTT	261
OY	247	AACGACGACATAGAAACCTGTGAAAACCTTGGACGTGATGTCATACAGTTCGACAC	306
Db	262	AACCAAAAGGATGAACACTGTGACAAATTTGAGTGTTTATGCATTAATATATGACAC	321
OY	307	ATTGGGTGAATATGAAAGTAAATTAATTAATCCCAAGCGTTGGCTGAGAGCTATA	366
Db	322	AGCTCTTGTCATTT-----ATTTAACTTTCTGCAAAACCTTTGGCTGCACAAACTGCA	375
OY	367	GACTATGCCACAACTCATCTCTGATTTGCTAGAAAC--CACCTTCTCTTGTTGGCTT	424
Db	376	GGGATATGGTAGAAACCAACTACGGATTTGCGCAAAACCAACACTCTCTCTTATATCT	435
OY	425	TTTATGTGGAACTGCTAGACAACTGTGAAACCT	459
Db	436	TTTTACTACAACTACAAAGCAATTTGTTGAAACCT	470

RESULT	13
LOCUS	AR207542
DEFINITION	AR207542 Sequence 535 bp
ACCESSION	AR207542
VERSION	AR207542.1
KEYWORDS	GI:21507322
SOURCE	.
ORGANISM	Unknown.
REFERENCE	Unknown.
AUTHORS	Unclassified.
TITLE	1 (bases 1 to 535)
JOURNAL	Colpitts,T.L.
FEATURES	Readends and methods useful for detecting diseases of the breast
source	Patent: US 6379671-A 1 30-Apr-2002;
	Location/Qualifiers
	1..535

BASE COUNT	153 a	127 c	105 g	150 t
ORIGIN	/organism="unknown"			
Query Match	47.0%	Score 223.8;	DB: 6;	Length 535;
Best Local Similarity	71.4%;	Pred. NO. 15e-55;		
Matches 329;	Conservative	0;	Mismatches 122;	Indels 8;
				Gaps 2;

QY	7	TGCGACGACGACTGAACACAGACAGCGCGGCTCGGCATGAGTGTGTAATGTC	66
Db	43	TGCGACCGCGAGTGAACACGACAGCAGCGCTCACCATGAAGTGTGTATGTC	102
QY	67	ATGCTGGCGGCTCTCTCTGCACTGCTATGCAATTCTGCTGCAAACTCTGGAGAC	126
Db	103	ATGCTGGCGGCTCTCTCCGACGACGCTACGAGGCTCTGCTCCCTTATGGAGAA	162
QY	127	ATGGTGAAGAAAGCATCAATTCGACATATCTATACCTGATATCAAGAGCTTTTCAA	186
Db	163	GTGATTTTCCAGAGCATCAATCCACAAGTGTCTAAGACTGATATCAAGAACTTTTCAA	222
QY	187	GAGTTCATGACAGTGATGCGGCTCGACAGGCTATGGGAAATTCGAAGTGTTCCTC	246
Db	223	GAGTTCATGACGACATGCGCTCACTCAAAATGCTATGATGAATGAAGGAATTTTCTT	282
QY	247	AACGAGCATATGAACTCTGAAAAACTTTGGACGATGAGTCAAGAGTATGCGACAG	306
Db	283	AACCAACGAGTATAAACTCTTGAGCATGTGGGTGTTTATGCAATTTATATGACAG	342
QY	307	ATTTGGTATATGAAGAGTAAATTACTTTACCCAAAGCGTGTGGCTCAGAGGGCTTAC	366
Db	343	AGCTTTGTGATTT-----ATTTTAACTTCTGCGAMAGACTTGTGCTCAGACAACTGCA	396
QY	367	GACTATGGCCAGAACTCATCTGTGTATTTGCTAGAAAC--CACTTCTTCTTGTGTGCTT	424
Db	397	GGGATGTGTGAACCACTACGGATGTGTCTCAACCAACCACTTCTTCTTTATATGCTC	456

```
OY      425 TTTATGTGGGAAC TGCTAGACA CTGTTGAAACT   459  
         ||| | |||| | |||| | |||| |  
Db       457 TTTTAC TCA AACTACAGA CACATT GTTG AA ACT   491
```

			RESULT 14	
			AX283024	
			LOCUS	
			DEFINITION	
			ACCESSION	
			VERSION	
			AX283024	
			Sequence	73 from Patent WO0175171.
			AX283024	
			AX283024.1	
			GI:16609951	
			DNA	linear
				PAT 02-NOV-2001

KEYWORDS
human,
Homo sapiens
SOURCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eultheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE	1
AUTHORS	Houghton, R. L., Dillon, D. C., Molesh, D. A., Xu, J., Zehentner, B. and Persing, D. H.
TITLE	Methods, compositions and kits for the detection and monitoring of breast cancer
JOURNAL	Patent: WO 0175171-A 73 11-OCT-2001; CORIAX CORPORATION (US)

FEATURE	Location	Qualifiers
SOURCE	1.	.503
BASE COUNT	/Organism="Homo sapiens"	
ORIGIN	/db_xref="taxon:9606"	
	145 a	119 c 97 g 142 t

Query Match: 46.78; Score 222.2; DB 6; Length 503;

Best Local Similarity /1.2%; Pred. NO. 4.3e-55;
Matches 324; Conservative 0; Mismatches 123; Indels 8; Gaps 2;

7 TGGCAGCGACGACTGACACACACAGCAGCGCCGCTGGCCATGAAGCTGCTTATGGTCTC 66

Db 22 TGGCACC CGGCACTGAACACCGACAGCAGACCTCACCATGAAGTTCTGTGATGGTCTC 81

67 ATGTGGGGCCCTCTCCTGCACTGCTATGCAGATTCTGGCTGCAAACTCTGTGAGGAC 126

Db 82 ATGCTGGGGCCCCCTCCACAGCACTGCTACGCAGGCTCTGGCTGCCCTTATTGGAGAT 141

127 ATGCTGAAGAAGACCATCAATTCGACATATCTATACCGAATACAAAGACTTCTCAA 186

D0: 167 CACGCGCTATGTACTCCAGTTCGGCCTGCATTGCCGAATAAAGAATTTGGCATTAACAATGATGATGCAAA
D8 142 GIGATTTCCAAGCACATCAATCCCACAGTGTCCTAACGACATGAATACAAAAGACTCTTTCAA 201

202 GAGTTCATAGAGCAGCAATGSCCACTACAAATGCGCATAGATGCAATTTGCAAGCAATCTGTTTTTTT 261
107 GGGATTCATAGAGCAGCAATGSCCACTACAAATGCGCATAGATGCAATTTGCAAGCAATCTGTTTTTTT 240

247 AACAGTCACATAGAACTCGAAAACTTTGCAGTGTGTCATACAGTGTACGACAGC 306

Db 262 AACCAACGGATGAAACTCTGACCAATGTTGAGCTGTTCTGCAATTAAATATATGACAGC 321

307 ATTTGGTGTATATGCAAGATTAATTACTTTACCCAGGGTTGGCTCAGAGGGCTACA 366

Db 322 AGCTTTTGATTT-----ATTTAACTTCTGCAAGACTTTGGCTCACAGAATGCA 375

367 GACTATGGCCAGAACTCATCTCTGTTGATTGCTAGAAC--CACTTCTCTCTTGTGTGCTT 424

Db	376	GGGTATVGGTCAGAAACCACTACGAGATTGCTGCAACACACACCCTTCTCTTCTTATCTCT	435
----	-----	---	-----

0y 425 TTTATGTGGGAAGCTGCTAGACAACTGTTGAACCT 459
||| | |||| | ||||| ||||| |||||

Db 436 TTTTACTACAACACTACAGACAATTGTTGAACCT 470

RESULT 15
1101111111

LOCUS	AX213264	429 bp	DNA	linear	PAT 06-SEP-2001
DEFINITION	Sequence 49 from Patent WO0158947				

REVISION	SEQUENCE	FROM	RECORD	NOV150257
ACCESSION	AX213264			
VERSION	AX213264.1	GI:15524179		

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 429)

AUTHORS

Carter, D., Vedvick, T.S., Valliave-Douglas, J., Houghton, R.L. and Dillon, D.C.

TITLE

Lipophilin complexes for use in cancer diagnosis and therapy

Patient: WO 0158947-A 49 16-AUG-2001;

CORIXA CORPORATION (US)

FEATURES

Location/Qualifiers

1..429

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT

121 a 106 c 87 g 115 t

ORIGIN

Query Match 44.5%; Score 211.6; DB 6; Length 429;

Best Local Similarity 72.5%; Pred. No. 6; Le-52;

Matches 290; Conservative 0; Mismatches 104; Indels 6; Gaps 1;

OY 7 TCCACGCGACGACTGACAGACAGACGCGCCTCGCATGAAGCTGCTGATGTCCTC 66
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 19 TGCACCGCGCGACTGACAGACGCGCGCTCCACGATGAAGTTGCTGATGTCCTC 78
OY 67 ATGCTGGCGCGCT 126
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 79 ATGCTGGCGCGCT 138
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 127 ATGCTGAAAGACCATCATATTCGACATATCTATACCTGAATACAAAGAGCTTCTCA 186
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 139 GTGATTTCCAGACATCAATCCACAAAGTGTCTAAGACTGAATACAAAGAGCTTCTCA 198
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 187 GAGTTCATGACAGCTAGTGGCGCGCTGACAGGCTATGGGAATTCAGCAGTGTCTC 246
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 199 GAGTTCATGACAGCTAGTGGCGCGCTGACAGGCTATGGGAATTCAGCAGTGTCTC 258
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 247 AACCACTCATAGACTCTGAAAACTTTGGACTGATGATGATGATGATGATGATGATG 306
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 259 AACCACTCATAGACTCTGAAAACTTTGGACTGATGATGATGATGATGATGATGATG 318
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
OY 307 ATTTGGTGAATATGAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 366
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 319 AGTCTTTGATTTT-----ATTTAACTTTTGGCAAGAGCTTTGGCTCAGAACTGCA 372
OY 367 GACTATGGCAGAGCTCATCTGTTGATTTGCTAGAAACAC 406
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 373 GGTATGTTGAGAAACAACTACGATTTGCTAGAAACAC 412
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

Search completed: January 25, 2003, 12:40:17
Job time : 3279 secs